



BIT-ERROR-RATE-MINIMIZING CHANNEL SHORTENING
USING POST-FEQ DIVERSITY COMBINING AND A GENETIC ALGORITHM

THESIS

Gökhan Altın, First Lieutenant, TUAF

AFIT/GE/ENG/09-01

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GE/ENG/09-01

BIT-ERROR-RATE-MINIMIZING CHANNEL SHORTENING
USING POST-FEQ DIVERSITY COMBINING AND A GENETIC
ALGORITHM

THESIS

Presented to the Faculty
Department of Electrical and Computer Engineering
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
In Partial Fulfillment of the Requirements for the
Degree of Master of Science in Electrical Engineering

Gökhan Altın, B.S.E.E.

First Lieutenant, TUAf

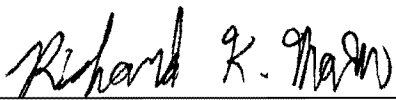
March 2009

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

BIT-ERROR-RATE-MINIMIZING CHANNEL SHORTENING
USING POST-FEQ DIVERSITY COMBINING AND A GENETIC
ALGORITHM

Gökhan Altın, B.S.E.E.
First Lieutenant, TUAF

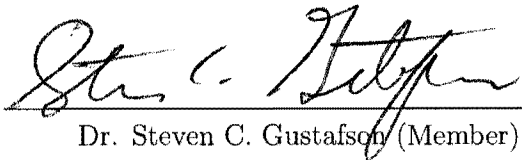
Approved:



Dr. Richard K. Martin (Chairman)

20 Mar 2009

date



Dr. Steven C. Gustafson (Member)

16 Mar 09

date



Maj Michael J. Mendenhall, PhD
(Member)

10 Mar 09

date

Abstract

In advanced wireline or wireless communication systems, i.e., DSL, IEEE 802.11a/g, HIPERLAN/2, etc., a cyclic prefix which is proportional to the channel impulse response is needed to append a multicarrier modulation (MCM) frame for operating the MCM accurately. This prefix is used to combat inter symbol interference (ISI). In some cases, the channel impulse response can be longer than the cyclic prefix (CP). One of the most useful techniques to mitigate this problem is reuse of a Channel Shortening Equalizer (CSE) as a linear preprocessor before the MCM receiver in order to shorten the effective channel length.

Channel shortening filter design is a widely examined topic in the literature. Most channel shortening equalizer proposals depend on perfect channel state information (CSI). However, this information may not be available in all situations. In cases where channel state information is not needed, blind adaptive equalization techniques are appropriate. In wireline communication systems (such as DMT), the CSE design is based on maximizing the bit rate, but in wireless systems (OFDM), there is a fixed bit loading algorithm, and the performance metric is Bit Error Rate (BER) minimization.

In this work, a CSE is developed for multicarrier and single-carrier cyclic prefixed (SCCP) systems which attempts to minimize the BER. To minimize the BER, a Genetic Algorithm (GA), which is an optimization method based on the principles of natural selection and genetics, is used.

If the CSI is shorter than the CP, the equalization can be done by a frequency domain equalizer (FEQ), which is a bank of complex scalars. However, in the literature the adaptive FEQ design has not been well examined. The second phase of this thesis focuses on different types of algorithms for adapting the FEQ and modi-

ifying the FEQ architecture to obtain a lower BER. Simulation results show that this modified architecture yields a 20 dB improvement in BER.

Acknowledgements

I would like to express my sincere appreciation to my faculty advisor, Dr. Richard K. Martin, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated. I also thank to Dr. Steven C. Gustaffson and Maj. Michael J. Mendenhall for their contribution to my thesis.

I would like to thank my family and my friend 1Lt. Adem Okal for their support and latitude provided to me in this endeavor.

I do also owe my great Turkish nation for providing me the chance to attend this education program.

Gökhan Altın

Table of Contents

	Page
Abstract	iv
Acknowledgements	vi
Table of Contents	vii
List of Figures	ix
List of Tables	x
List of Symbols	xi
List of Abbreviations	xii
 I. Introduction	 1
1.1 Multicarrier Modulation	1
1.1.1 Discrete Multitone	2
1.1.2 Orthogonal Frequency Division Multiplexing . .	4
1.2 Single-Carrier Cyclic Prefixed Modulation	5
1.3 Channel Shortening	7
 II. Literature Survey	 11
2.1 Channel Shortening Equalizer Design Methods	11
2.1.1 Minimum Mean Squared Error Method	11
2.1.2 Maximum Shortening Signal to Noise Ratio Method	14
2.1.3 Maximum Geometric Signal to Noise Ratio Method	17
2.2 Adaptive Channel Shorteners	19
2.2.1 Multicarrier Equalization by Restoration of Redundancy	21
2.2.2 Sum-Squared Auto-correlation Minimization . .	22
2.2.3 Single Lag Auto-correlation Minimization	24
2.2.4 Blind Adaptive Channel Shortening Equalizer Algorithm which can Provide Shortened Channel State Information	25
2.3 Genetic Algorithm	27
2.3.1 Selecting the Variables and the Cost Function .	28
2.3.2 Initial Population	28
2.3.3 Natural Selection	28
2.3.4 Mating	30

	Page
2.3.5 Mutation	30
2.3.6 The Next Generation and Convergence	31
III. Bit-Error-Rate-Minimizing Channel Shortening using Post-FEQ Diversity Combining and a Genetic Algorithm	32
3.1 BER Minimizing Channel Shortening using Post-FEQ Diversity Combining	33
3.1.1 Review of the RLS-MERRY Algorithm	36
3.1.2 Adaptive FEQ	39
3.2 BER Minimizing Channel Shortening with a Genetic Algorithm	47
3.2.1 System Model	48
3.2.2 BER Models	49
3.2.3 Applying Genetic Algorithm	54
IV. Simulations and Results	57
4.1 Results for Post-FEQ Diversity Combining	57
4.2 Results for Channel Shortening with GA	59
4.2.1 Results of GA for Multicarrier Systems	59
4.2.2 Results of GA for SCCP Systems	62
V. Conclusions	68
Bibliography	71
Vita	75

List of Figures

Figure		Page
1.1.	MIMO OFDM System model.	4
1.2.	MIMO SCCP System model.	6
1.3.	Channel shortening.	7
2.1.	The MMSE equalizer block diagram.	12
2.2.	The MSSNR design idea.	16
2.3.	Blind signal processing methods.	20
2.4.	Difference in the ISI at the received CP and at the end of the received symbol	21
2.5.	System model for SAM.	23
2.6.	Flowchart of a GA.	29
2.7.	Two parents mate to generate two children. Then, these two children are added to the population.	31
3.1.	SIMO multicarrier system model.	33
3.2.	Post-FEQ diversity combining with 2 FFT.	34
3.3.	Post-FEQ diversity combining with individual FFT.	34
3.4.	An example of a fading channel.	40
4.1.	MERRY Cost vs. CSE iteration	58
4.2.	BER vs. FEQ iteration after the CSE has converged	59
4.3.	BER vs. SNR for a multicarrier system	61
4.4.	SER vs. SNR for a multicarrier system	62
4.5.	BER vs. SNR for SCCP system	64
4.6.	SER vs. SNR for aSCCP system	65
4.7.	BER vs. SNR for SCCP system	66
4.8.	SER vs. SNR for SCCP system	67

List of Tables

Table		Page
1.1.	Example values of OFDM System Parameters.	5
3.1.	Rank weighting in a GA	55
3.2.	An example of mutation and new population	56

List of Symbols

Symbol		Page
L	Number of Transmit Antenna	9
P	Number of Receive Antenna	9
N	(I)FFT Length	9
v	Cyclic Prefix Length	9
i	Subchannel Index	9
Δ	Delay	9
\mathbf{x}	Received Signal	9
\mathbf{h}	Channel	9
$\mathbf{c} = \mathbf{w} \star \mathbf{h}$	Shortened Channel	9
\mathbf{r}	Received Signal	9
\mathbf{y}	Output Signal	9
$(\cdot) \star (\cdot)$	Convolution	9
$(\cdot)^*$	Complex Conjugate	10
$(\cdot)^T$	Matrix Transpose	10
$(\cdot)^H$	Conjugate Transpose	10
$\mathcal{E}\{\cdot\}$	Statistical Expectation	10
N_{var}	Variable Number	28
N_{gene}	Gene Number	28
ρ	Forgetting Factor	38
μ	Step Size	40
$e(n)$	Error	40
γ	Dispersion Constant	44
\mathbf{D}	FEQ Vector	46
\odot	Element-by-Element Multiplication	47
N_a	Active Tone Numbers	48
\mathcal{F}	FFT Operation	49

List of Abbreviations

Abbreviation		Page
DSL	Digital Subscriber Line	1
ISI	Inter-Symbol Interference	1
MCM	Multicarrier Modulation	1
DFE	Decision Feedback Equalizer	1
DSP	Digital Signal Processor	2
FFT	Fast Fourier Transform	2
DMT	Discrete Multitone	2
OFDM	Orthogonal Frequency Division Multiplexing	2
QAM	Quadrature Amplitude Modulation	2
DFT	Discrete Fourier Transform	3
ICI	Inter-Carrier Interference	3
CP	Cyclic Prefix	3
FIR	Finite Impulse Response	3
TEQ	Time Domain Equalizer	4
FEQ	Frequency Domain Equalizer	4
FEC	Forward Error Correction	5
MIMO	Multiple Input Multiple Output	5
SCCP	Single-Carrier Cyclic Prefixed	5
SC-FDE	Single-Carrier Frequency Domain Equalization	5
BER	Bit Error Rate	8
DD-LMS	Decision Directed Least Mean Squared	9
DD-RLS	Decision Directed Recursive Least Squares	9
MMSE	Minimum Mean Squared Error	11
ML	Maximum Likelihood	11
TIR	Target Impulse Response	11

Abbreviation		Page
UTC	Unit Tab Constraint	12
UEC	Unit Energy Constraint	12
MSSNR	Maximum Shortening Signal to Noise Ratio	14
Min-ISI	Minimum Inter-Symbol Interference	16
Min-IBI	Minimum Inter-Block Interference	17
MDS	Minimum Delay Spread	17
AWGN	Additive White Gaussian Noise	17
MGSNR	Maximum Geometric Signal to Noise Ratio	18
MERRY	Multicarrier Equalization by Restoration of Redundancy .	21
SAM	Sum-Squared Auto-correlation Minimization	22
SLAM	Single Lag Auto-correlation Minimization	24
BACS-SI	Blind Adaptive Channel Shortening Equalizer Algorithm with State Information	25
GA	Genetic Algorithm	27
CSE	Channel Shortening Equalizer	32
SIMO	Single-Input Multiple-Output	33
LMS	Least Mean Square	40
RLS	Recursive Least Squares	41
CMA	Constant Modulus Algorithm	44
IR	Iteratively Reweighted	47
GN	Gauss-Newton	47
MER	Minimum Error Rate	47

BIT-ERROR-RATE-MINIMIZING CHANNEL SHORTENING USING POST-FEQ DIVERSITY COMBINING AND A GENETIC ALGORITHM

I. Introduction

MULTIMEDIA applications and the development of the Internet have led to the need for high speed and wide band digital communications. Development of these technologies and the need to get information in very short times shows that the demand for wide band technologies will increase.

This demand for high speed communication induces service providers to provide high-quality low-cost systems. Thus, communication systems have improved from voice band applications to Digital Subscriber Lines (DSL) and wireless networking [1].

1.1 *Multicarrier Modulation*

High speed communications require broad band channels, in which Inter-Symbol Interference (ISI) is one of the main problems. This undesirable effect causes neighboring symbols to interfere with each other at the receiver. In frequency domain, this effect varies the magnitude and phase response of the channel. If ISI is not inhibited, it causes information losses. Channel equalization and Multicarrier Modulation (MCM) are two approaches to combat ISI [2].

In channel equalization, an equalizer filter is used to undo the spectral shaping effect of the channel. For this purpose, two approaches, i.e., linear and nonlinear equalizer structures, are considered in the literature. Although nonlinear equalizers, e.g., Decision Feedback Equalizers (DFE), give better results in channel equalization, they have high computational complexity, especially under high sampling rates. Therefore, linear equalizers, which are relatively simpler than the others, have gained more popularity [3, 4].

In narrow band systems, the channel can be considered flat and the ISI will not occur. Additionally, the communication speed can be reduced as well to decrease occurrence of the ISI. In this case, the balance between the wide band, high speed with-ISI systems and narrow band, low speed without-ISI systems can be made by MCM. In MCM, the channel is partitioned into small bandwidth channels called subchannels. If a subchannel is narrow enough, then the subchannels are independent from each other and the ISI would not occur in these low speed subchannels. Thus, information can be transmitted in these subchannels without ISI and the total number of bits is the sum of the bits transmitted over these subchannels.

Reasons for MCM popularity are [1]:

- MCM performs channel equalization easily, by scalars rather than a filter,
- It controls the power and the number of bits in the subchannels and thus uses the bandwidth efficiently,
- Due to its long symbol duration, it is resistant to impulsive noise and fading,
- Advancements in Digital Signal Processors (DSP) have made it very easy to implement MCM.

The Fast Fourier Transform (FFT) is one of the efficient ways to divide a channel into the numerous of subchannels. If an FFT is used to implement multicarrier modulation, it is called Discrete Multitone (DMT) or Orthogonal Frequency Division Multiplexing (OFDM). DMT is generally used for wireline systems and OFDM is used for wireless systems. The main differences between these applications are assignment of the power and loading the bits into subchannels.

1.1.1 Discrete Multitone. Discrete multitone is an MCM method in which the communication channel is divided into a large number of subchannels via an FFT. The input bit stream in the transmitter is buffered and mapped into a $N/2 \times 1$ complex vector \mathbf{X}_i at time i using a modulation method, such as Quadrature Amplitude Modulation (QAM) in ADSL. The process of mapping, partitioning the bit stream

into complex values, and assigning it to the subchannels according to SNR, such that high SNR subchannels receive more bits than low SNR ones, is called bit loading. The complex-valued symbols can be considered in the frequency domain, but complex valued symbols are modulated to time domain. Using $N/2$ independent modulators is very expensive, and instead, modulation can be done via Discrete Fourier Transform (DFT). DFT can be implemented efficiently by an N -point FFT.

In order to get time domain samples, the $\tilde{\mathbf{X}}_i^*$ which is the mirrored and conjugate symmetric version of \mathbf{X}_i is grouped with \mathbf{X}_i to obtain a $N \times 1$ length, $\hat{\mathbf{X}}_i = [\mathbf{X}_i^T \tilde{\mathbf{X}}_i^{*T}]^T$ vector, then this vector is sent to the input of the inverse-FFT (IFFT) block. The $\hat{\mathbf{X}}_i$ vector which is ready to send over a telephone line at the end of the IFFT has *real valued* DMT symbols.

As is well known, convolution in time domain is equivalent to multiplication in frequency domain after Fourier transform. The same relation is true for the DFT, but as a circular convolution. In a communication system, due to the fact that sending the data symbols over a channel causes linear convolution, it is not quite a multiplication of received signals in the frequency domain after the DFT. To address this problem, the transmitted signal should be converted to a circular form.

If the length of the channel impulse response is less than or equal to $v + 1$, ISI is completely eliminated by adding a guard period of length v . Additionally, if this guard period is chosen as the copy of the last v samples of a DMT frame, the signal is circular and Inter-Carrier Interference (ICI) does not occur. This type of guard period is called Cyclic Prefix (CP). The group of $N + v$ samples in time domain is named a DMT frame, and frames are sent through the channel sequentially.

The CP of v samples reduces the data rate by a factor of $N/(N + v)$. In order to eliminate the ISI, a long guard period should be added in a relatively long channel, but this long guard period reduces the data rate. To solve this dilemma, an equalizer which shortens the channel impulse response to the length v or less is added to the receiver as a processor. This Finite Impulse Response (FIR) filter, called

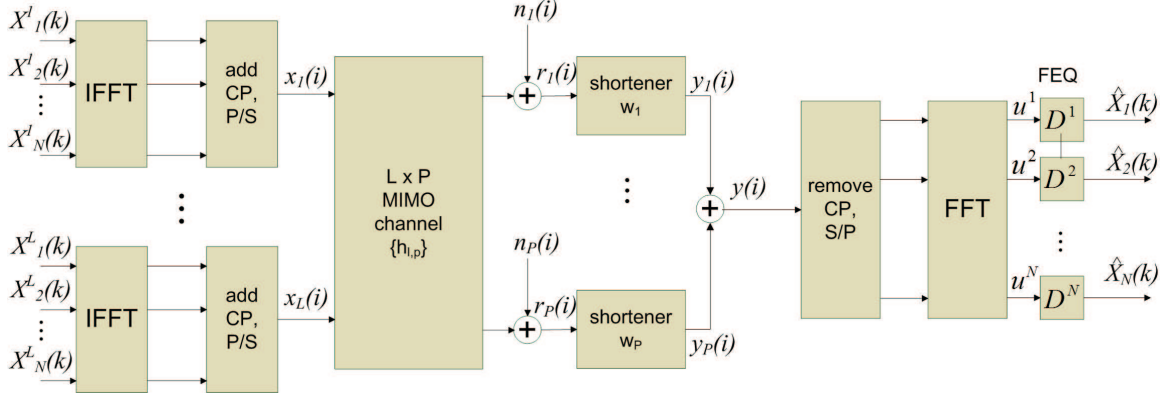


Figure 1.1: MIMO OFDM system model [12].

the Time Domain Equalizer (TEQ), shortens the equalized channel impulse response to the predetermined CP length or less. Although TEQ shortens the channel in the time domain, phase deviation and magnitude attenuation are still present in each subchannel in the frequency domain. These variations can be removed by using a Frequency Domain Equalizer (FEQ) [1, 5, 6].

1.1.2 Orthogonal Frequency Division Multiplexing. Orthogonal Frequency Division Multiplexing (OFDM) is an MCM method similar to DMT. OFDM is popular for wireless systems such as IEEE 802.11a/g wireless LAN, HIPERLAN/2, MMAC [7], wireless metropolitan area networks (IEEE 802.16) [8], digital video and audio broadcasting in Europe [9, 10], satellite radio (e.g., Sirius and XM Radio) [11], and it is the proposed standard for multiband ultrawideband (IEEE 802.15.3a). OFDM is a frequency division multiplexing method in which the available bandwidth is divided into numerous nearly orthogonal subchannels. As in DMT, the modulation in OFDM can be done using FFT/IFFT pairs. An OFDM symbol includes the data and the cyclic prefix, and ISI/ICI can be eliminated using the cyclic prefix similarly in DMT. Bit loading to the subchannels is based on the SNR in DMT, but the number of bits in each subchannel is equal and constant in OFDM [5].

As a wireless application, OFDM faces multipath delay spread; therefore the received symbols are the summation of the replicas of the transmitted signal. If the

Table 1.1: Example values of OFDM System Parameters for IEEE 802.11a [13].

Parameters	Value
Ratified Year	1999
Frequency Range	4615-5825 MHz
Channel Spacing B	20 MHz
Number of Subcarriers N	52
Modulation	BPSK, QPSK, 16-QAM, 64-QAM
Useful Symbol Length T_U	$3.2 \mu s$
Additional Guard Interval T_G	$0.8 \mu s$
Subchannel Spacing	312.5 kHz
Data Rate	6-54 Mbps
Link Spectral Efficiency R/B	0.30-2.7 bit/s/Hz
Inner FEC	Conv Coding with code rates $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}$

CP is longer than the delay spread, ISI/ICI does not occur. In some certain frequency bands, the multipath fading channels have deep nulls. These nulls force the OFDM subchannels to be unable to carry data. One way to mitigate this undesirable effect is Forward Error Correction (FEC) technique. Another way is frequency domain interleaving, which spreads the information into the various subchannels to increase the probability of correct data decoding. An OFDM receiver performs time domain equalization to reduce the channel length, to simplify decoding, and to reduce the delay spread [5].

A block diagram of an Multiple Input Multiple Output (MIMO) OFDM system is shown in Fig.1.1. The parameters standardized for the OFDM (IEEE 802.11a) are in Table 1.1.

1.2 Single-Carrier Cyclic Prefixed Modulation

Single-Carrier Cyclic Prefixed (SCCP) modulation is also known as Single-Carrier Frequency Domain Equalization (SC-FDE). SCCP modulation has not been studied extensively, but it is gaining importance in the literature.

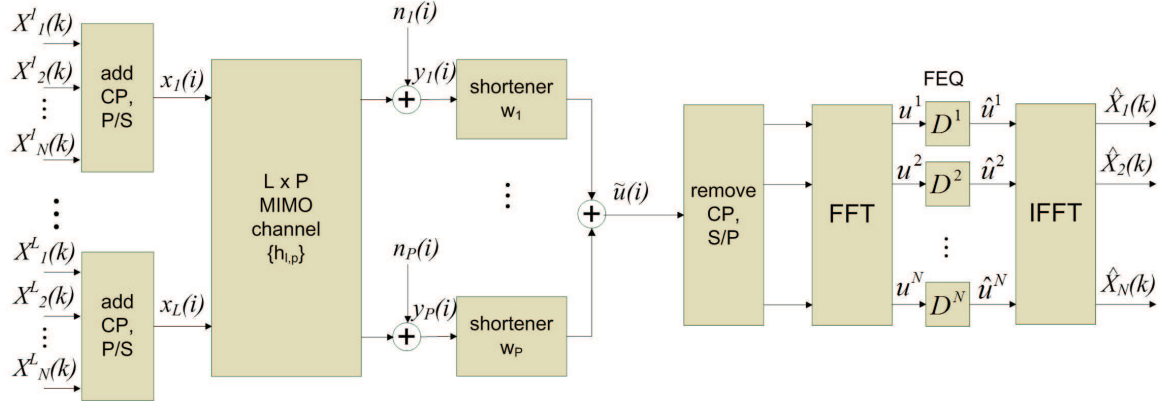


Figure 1.2: MIMO SCCP system model [12].

A SCCP system model is shown in Fig. 1.2. The concept of SCCP looks similar to that of OFDM, but the IFFT is removed from the transmitter and placed at the receiver, after the FEQ. As a result, there are no null tones as in OFDM, due to the fact that the transmitter operates completely in the time domain.

Moreover, OFDM is not very effective in deep fading channels due to the deep nulls, and it requires error correction coding with frequency domain interleaving. In contrast, SCCP systems can work without error correction coding [14].

The use of SC modulation and Frequency Domain Equalization (FDE) by processing the FFT of the received signal has several attractive features:

- SC modulation has reduced peak to average ratio requirements from OFDM, thereby allowing the use of less costly power amplifiers,
- Its performance with FDE is similar to that of OFDM, even for very long channel delay spread,
- Frequency domain receiver processing has a similar complexity reduction advantage to that of OFDM: complexity is proportional to log of multipath spread,
- Coding, while desirable, is not necessary for combatting frequency selectivity, as it is in nonadaptive OFDM,
- SC modulation is a well-proven technology in many existing wireless and wireline applications, and its RF system linearity requirements are well known [15].

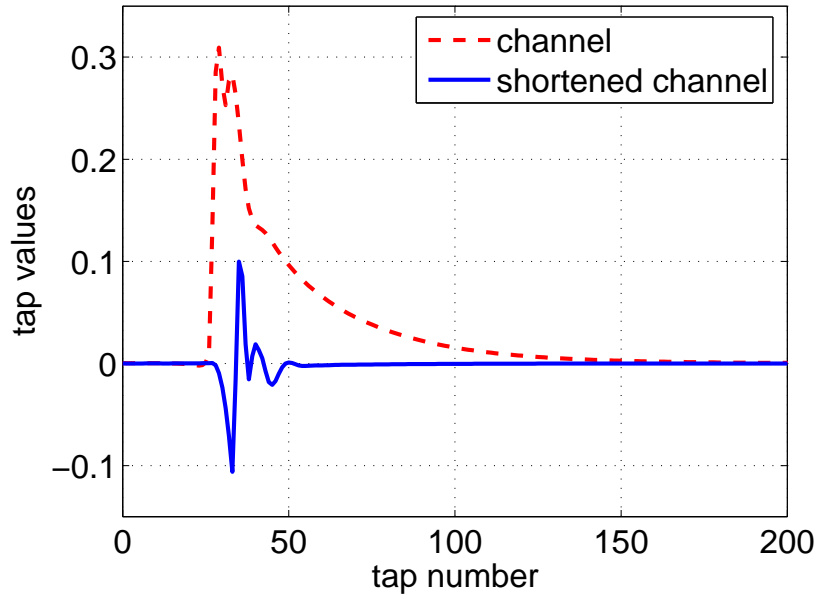


Figure 1.3: Channel shortening.

1.3 Channel Shortening

The goal of the full channel equalization is to prevent the corruptive effect of the ISI channel with an equalizer. In other words, convolution of the channel and the equalizer is converted to a Kronocker delta function ($\delta(\Delta)$), i.e., an impulse function [16] (Here, Δ is a convenient delay and is a design parameter.).

In a multicarrier system, there is no need to equalize the channel fully. Instead, shortening the channel to a *predetermined length* by designing an equalizer as a processor in the receiver is more suitable for system efficiency. Designing an equalizer with this goal is called *channel shortening*. The channel shortening idea is shown in Fig. 1.3.

Channel shortening first appeared in 1970s to reduce the complexity of Viterbi algorithm [17], then in 1990s it became a current issue with multicarrier modulation [18]. Channel shortening in recent years came into use for multiuser detection [19]. CSEs are used for computational complexity reduction for multiuser detection by blocking the signals from a subset of users and detecting the remainder. Channel

shortening is also used to reduce the complexity of ultra wideband systems. A pulse is reduced by shortening the multipath channel, reducing the number of correlators needed ultra wideband systems [20]. Also, as in acoustics, a channel shortener can be used to optimize the D50-measure, defined by psychoacoustics, for the intelligibility of speech as the ratio of energy in a 50-ms window of the room impulse response to the total energy of the impulse response [21].

The channel impulse response should be known in the channel shortening system models; therefore a training sequence is sent with the transmitted symbol. However, in some cases (e.g., DVB), the receiver can not fully receive the training sequence in the time domain [22]. On the other hand, to send this training sequence with information bearing signals reduces the channel capacity. For these reasons, blind, adaptive channel shortening techniques without training and channel knowledge have been proposed.

Early channel shortening techniques were based on heuristic objective functions. In recent algorithms, the goal is maximizing the bit rate for a given Bit Error Rate (BER), which is a proper performance measurement in wireline multicarrier systems, i.e., DSL [23–25]. In contrast, wireless multicarrier, i.e., OFDM, and SCCP systems have a fixed bit loading algorithm, and the performance metric is the BER for a fixed bit rate. Additionally, the CSE can be updated to minimize the BER for the initial bit loading even in DSL. Here, in Chapter 3, a BER minimizing CSE for multicarrier and SCCP systems is investigated.

In Chapter 2, early channel shortening equalizer design algorithms are reviewed. The Minimum Mean Squared Error (MMSE) method [17], the Maximum Shortening Signal to Noise Ratio (MSSNR) method [26], and the Maximum Geometric Signal to Noise Ratio (MGSNR) [27] method are summarized. Also, adaptive channel shorteners, Multicarrier Equalization by Restoration of Redundancy (MERRY) [28], Sum-Squared Auto-correlation Minimization (SAM) [29], Single Lag Auto-correlation Minimization (SLAM) [30], and Blind Adaptive Channel Shortening Equalizer Algorithm

which can provide Shortened Channel State Information (BACS-SI) [31] methods, are reviewed. Finally, the Genetic Algorithm (GA) which is an optimization method based on the principles of natural selection and genetics, is described.

Chapter 3 is separated into two parts. As stated earlier, the channel delay spread must be shorter than the CP. If the channel can be shortened with an CSE, the equalization can be done by an FEQ, which is just a bank of complex scalars. The first part of Chapter 3 proposes to modify an OFDM system architecture to obtain a lower BER. Three types of adaptation rules, i.e., Decision Directed Least Mean Squared (DD-LMS), DD-Recursive Least Squares (DD-RLS), and Constant Modulus Algorithm (CMA), are derived for the new architectures and are used to adapt FEQs. To provide the minimum BER, different types of diversity combining are considered. The second part of chapter 3 proposes to design a CSE adapted with a Genetic Algorithm (GA) to obtain minimum BER.

Chapter 4 provides the simulation results. In simulations, the BER performance of DD-LMS, DD-LMS and CMA algorithms is compared with the Adaptive Minimum Mean Squared Algorithm (A-MMSE) [17] for reference. The BER performance for GA is presented as well. Simulation results show that the proposed post-FEQ combining technique has a two order of magnitude improvement, i.e., 20 dB, in BER. Chapter 5 concludes the thesis.

The notation used in this thesis are as follows:

- We assume MIMO channel model with L transmit antennas and P receive antennas.
- N is the (I)FFT length, v CP length, i subchannel index and Δ delay.
- Vectors are bold, lower case; matrices are bold, upper case; and \mathbf{x} , \mathbf{h} , $\mathbf{c} = \mathbf{w} \star \mathbf{h}$, \mathbf{r} and \mathbf{y} are CSE with length L_w , channel with length L_h , shortened channel with length $L_c = L_w + L_h - 1$, received signal and the output of a CSE, respectively. Here, $(\cdot) \star (\cdot)$ denotes convolution.

- $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote complex conjugate, matrix transpose and conjugate transpose, respectively, and $\mathcal{E}\{\cdot\}$ denotes statistical expectation.

Part of this thesis has been presented at the 42nd Asilomar Conference on Signals, Systems and Computers, Asilomar Hotel and Conference Ground, Pacific Grove, CA, October 26th-29th, 2008 and has been submitted to IEEE Communications Letters. Also, the thesis will be submitted to IEEE Transactions on Signal Processing.

II. Literature Survey

2.1 Channel Shortening Equalizer Design Methods

Because of its simplicity, the Minimum Mean Squared Error (MMSE) method is a well studied method in the literature [17]. The MMSE minimizes the error of the equalizer after the channel and a Target Impulse Response (TIR), which is a design parameter. Also, Adaptive MMSE design methods are commonly used in practical systems. The Maximum Shortening Signal to Noise Ratio (MSSNR) method [26] minimizes the energy of the part of the channel impulse response that causes Inter Symbol Interference (ISI). Both the MMSE and the MSSNR do not attempt to maximize channel capacity. Al-Dhahir and Cioffi [27] propose a method called Maximum Geometric SNR (MGSNR) to shorten the channel impulse response while maximizing the bit rate.

Other methods proposed in the literature are Min-ISI [23], Min-IBI [32], Maximum Bit Rate (MBR) [23], Dual Path and Per Tone Equalizers (PTE) [33].

2.1.1 Minimum Mean Squared Error Method. The computational complexity of the Viterbi algorithm is exponential in the length of channel impulse response. This complexity can be reduced by a processor in the receiver. Falconer and Magee [17] propose the MMSE method to shorten the channel impulse response for Maximum Likelihood (ML) receivers. After Falconer and Magee, Chow and Cioffi apply this method to MCM [34].

The design idea of the MMSE method is shown in Fig. 2.1. The system architecture consists of two Finite Impulse Response (FIR) filters. One is in the upper branch after the channel and the other is in the lower branch as a Target Impulse Response (TIR). The convolution of the channel and the filter in the upper branch is called the *equalized channel impulse response*. The lower branch is only for design concern and is a *virtual* branch. The goal of the method is to find the equalizer (Time Domain Equalizer, (TEQ)) coefficients which minimize the error between the signals of the output of the equalizer and the output of the TIR.

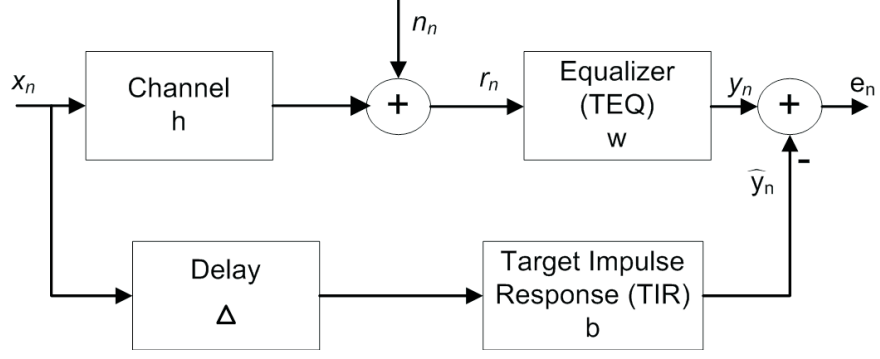


Figure 2.1: The MMSE equalizer block diagram.

In order to understand the idea behind the MMSE method, assume that the error is zero for an input signal. In this case, the impulse responses of both branches are equal. Thus, the equalized channel impulse response (upper branch) is equal to the delayed version of the TIR. In other words, if the coefficients of the TIR are chosen as (predetermined) *fixed length*, the channel impulse response is forced to the *same length*.

Setting the taps of equalized channel impulse response and the TIR to zero forces the error to zero. In this case, information can not be received at the receiver. One of the proposed solutions is to impose a unit tap constraint (UTC) [34]. A better solution is a unit energy constraint (UEC) that the norm of the TIR is a constant (usually 1) [35, 36].

The FIR output of the channel is

$$\mathbf{r}_n = \mathbf{H}\mathbf{x}_n + \mathbf{n}_n. \quad (2.1)$$

Here, \mathbf{H} is the channel convolution matrix, \mathbf{x}_n is the transmitted signal, and \mathbf{n}_n is noise. The output of the n_w length equalizer, $\mathbf{w} = [w_0 \ w_1 \ \dots \ w_{n_w-1}]^T$ is

$$y(n) = \mathbf{w}^T \mathbf{r}_n = \mathbf{w}^T \mathbf{H}\mathbf{x}_n + \mathbf{w}^T \mathbf{n}_n. \quad (2.2)$$

Here

$$\begin{bmatrix} r_{n_w-1} \\ r_{n_w-2} \\ \vdots \\ r_n \end{bmatrix} = \begin{bmatrix} h_0 & \cdots & h_{n_h} & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{n_h} & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & \cdots & h_{n_h} \end{bmatrix} \begin{bmatrix} x_{n_h+n_w-1} \\ x_{n_h+n_w-2} \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} n_{n_w-1} \\ n_{n_w-2} \\ \vdots \\ n_n \end{bmatrix} \quad (2.3)$$

is the received signal. Assuming the n_b length TIR as $\mathbf{b} = [b_0 \ b_1 \ \cdots \ b_{n_b-1}]^T$, and augmented TIR with Δ as $\tilde{\mathbf{b}} = [\mathbf{0}_{1 \times \Delta} \ \mathbf{b}^T \ \mathbf{0}_{1 \times (n_h+n_w-1-n_b-\Delta)}]^T$, ($\mathbf{0}$ is the matrix of zeros), the output of the lower branch at time n is

$$\hat{y}(n) = \tilde{\mathbf{b}}^T \mathbf{x}_n. \quad (2.4)$$

Using equations (2.1) and (2.4) and assuming that the signal and the noise are independent, the mean squared error (MSE) between the output of the equalizer and the TIR is

$$\begin{aligned} MSE(\mathbf{w}, \mathbf{b}) &= E \{ \epsilon(n)^2 \} = E \{ (y(n) - \hat{y}(n))^2 \} \\ &= E \left\{ \left((\mathbf{w}^T \mathbf{H} \mathbf{x}_n + \mathbf{w}^T \mathbf{n}_n) - \tilde{\mathbf{b}}^T \mathbf{x}_n \right)^2 \right\} \\ &= [\tilde{\mathbf{b}}^T - \mathbf{w}^T \mathbf{H}] \mathbf{R}_{xx} [\tilde{\mathbf{b}} - \mathbf{H}^T \mathbf{w}] + \mathbf{w}^T \mathbf{R}_{nn} \mathbf{w}. \end{aligned} \quad (2.5)$$

To find optimum equalizer taps, the gradient of MSE with respect to \mathbf{w} is set to zero:

$$\mathbf{w} = [\mathbf{H} \mathbf{R}_{xx} \mathbf{H}^T + \mathbf{R}_{nn}]^{-1} \mathbf{H} \mathbf{R}_{xx} \tilde{\mathbf{b}}. \quad (2.6)$$

By substituting (2.6) into (2.5),

$$MSE(\mathbf{b}) = \mathbf{b}^T \mathbf{R}_{\Delta} \mathbf{b}, \quad (2.7)$$

where $\mathbf{R}_{\Delta} = \mathbf{\Gamma}^T [\mathbf{R}_{xx}^{-1} + \mathbf{H}^T \mathbf{R}_{nn}^{-1} \mathbf{H}]^{-1} \mathbf{\Gamma}$, in which \mathbf{I}_{n_b+1} is the $n_b + 1$ identity matrix and $\mathbf{\Gamma} = [\mathbf{0}_{(n_b+1) \times \Delta} \ \mathbf{I}_{n_b+1} \ \mathbf{0}_{(n_b+1) \times (n_w+n_h-\Delta-n_b-1)}]^T$.

Equation (2.7) is minimized for $\mathbf{b} = \mathbf{0}$, but some constraints should be applied against this trivial solution.

2.1.1.1 Unit Tap Constraint (UTC). Al-Dhahir and Cioffi [35] define $\mathbf{e}_i = [\mathbf{0}_{1 \times (i-1)} \quad 1 \quad \mathbf{0}_{1 \times (n_b-1-i)}]$ as the i^{th} unit vector to obtain this solution. Using the Lagrangian, one can find i which is a value that minimizes the MSE as

$$i_{opt} = \arg \max_{0 \leq i \leq v} \{\mathbf{R}_{\Delta}^{-1}(i, i)\} \quad (2.8)$$

where $i \in [0, v]$ and $\mathbf{R}_{\Delta}^{-1}(i, i)$ is the i^{th} element of the diagonal matrix \mathbf{R}_{Δ}^{-1} . The optimum \mathbf{b} is

$$\mathbf{b}_{opt} = \frac{\mathbf{R}_{\Delta}^{-1} \mathbf{e}_{i_{opt}}}{\mathbf{R}_{\Delta}^{-1}(i_{opt}, i_{opt})}. \quad (2.9)$$

\mathbf{w}_{opt} can be found to substitute $\mathbf{b} = \mathbf{b}_{opt}$ in (2.6).

2.1.1.2 Unit Energy Constraint (UEC). Instead of UTC, the unit energy constraint, $\mathbf{b}^T \mathbf{b} = 1$ can be used. Again using the Lagrangian and taking the gradient with respect to \mathbf{b} and setting it to zero yields the solution which is the eigenvector corresponding to the minimum eigenvalue of \mathbf{R}_{Δ} . Here, the minimum value of the MSE, $MSE = \mathbf{b}^T \mathbf{R}_{\Delta} \mathbf{b} = \mathbf{b}^T \lambda \mathbf{b} = \lambda$, can be thought as the minimum eigenvalue of \mathbf{R}_{Δ} . Therefore, TIR, \mathbf{b} , should be chosen as the eigenvector corresponding to this eigenvalue.

2.1.2 Maximum Shortening Signal to Noise Ratio Method. Melsa, Younce and Rohrs [26] propose the Maximum Shortening Signal to Ratio (MSSNR) method to find a TEQ which keeps the energy inside a target window constant and minimizes the energy of the Shortened Impulse Response (SIR) outside the target window.

The signal is defined as the portion of the channel within the window:

$$\mathbf{h}_{win} = \begin{bmatrix} h_{\Delta} & h_{\Delta-1} & \cdots & h_{\Delta-n_w+1} \\ h_{\Delta+1} & h_{\Delta+2} & \cdots & h_{\Delta-n_w+2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{\Delta+v} & h_{\Delta+v-1} & \cdots & h_{\Delta-n_w+v+1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{n_w-1} \end{bmatrix} = \mathbf{H}_{win} \mathbf{w}, \quad (2.10)$$

and the portion outside the window is

$$\mathbf{h}_{wall} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ h_{\Delta-1} & h_{\Delta-2} & \cdots & h_{\Delta-n_w} \\ h_{\Delta+v+1} & h_{\Delta+v} & \cdots & h_{\Delta-n_w+v+2} \\ \vdots & \vdots & & \vdots \\ h_{n_c-1} & h_{n_c-2} & \cdots & h_{n_c-n_w+1} \\ 0 & h_{n_c-1} & \cdots & h_{n_c-n_w+2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & h_{n_c-1} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{n_w-1} \end{bmatrix} = \mathbf{H}_{wall} \mathbf{w} \quad (2.11)$$

The energy inside and outside the window is

$$\begin{aligned} \mathbf{h}_{win}^T \mathbf{h}_{win} &= \mathbf{w}^T \mathbf{H}_{win}^T \mathbf{H}_{win} \mathbf{w} = \mathbf{w}^T \mathbf{B} \mathbf{w} \\ \mathbf{h}_{wall}^T \mathbf{h}_{wall} &= \mathbf{w}^T \mathbf{H}_{wall}^T \mathbf{H}_{wall} \mathbf{w} = \mathbf{w}^T \mathbf{A} \mathbf{w}. \end{aligned} \quad (2.12)$$

The SSNR is

$$SSNR = \frac{\mathbf{w}^T \mathbf{B} \mathbf{w}}{\mathbf{w}^T \mathbf{A} \mathbf{w}}. \quad (2.13)$$

MSSNR minimizes the energy of the denominator of SSNR while keeping the energy of the numerator of SSNR equal to 1. The solution (details can be found

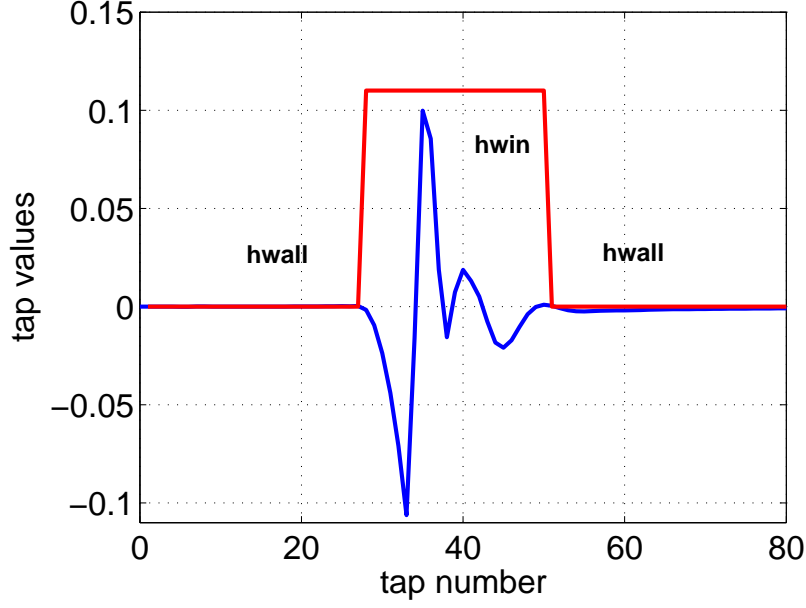


Figure 2.2: The MSSNR design idea.

in [26]) is

$$\mathbf{w}_{opt} = \left(\sqrt{\mathbf{B}} \right)^{-1} \mathbf{z}_{min} \quad (2.14)$$

where $\sqrt{\mathbf{B}}$ is the Cholesky decomposition of \mathbf{B} and \mathbf{z}_{min} is the eigenvector corresponding to the minimum eigenvalue of $(\sqrt{\mathbf{B}})^{-1} \mathbf{A} (\sqrt{\mathbf{B}^T})^{-1}$.

MSSNR minimizes the portion of the SIR that causes ISI. If the energy outside the target window is zero, the channel is shortened and the ISI is totally eliminated. Being zero of the outside of the window is the optimum solution and with this solution, ISI can be eliminated and the maximum channel capacity can be obtained. In practice this solution can not be achieved. For this reason, and because noise is ignored, the MSSNR is not guaranteed to maximize the channel capacity.

2.1.2.1 Extensions of the MSSNR Method. The MSSNR method has many extensions. The Minimum Inter-Symbol Interference (Min-ISI) evaluates the residual energy in the tails of the channel in the frequency domain and aims to place the extra energy in unused frequency bins. In this case, the \mathbf{A} and \mathbf{B} matrices are more complicated. Details of this method can be found in [23].

The Minimum Inter-Block Interference (Min-IBI) method [32] shapes the inter-block interference power, which increases linearly with the distance of the channel taps from the boundaries of the length $v + 1$ desired non-zero window. The \mathbf{A} and \mathbf{B} matrices are similar to MSSNR except with a linear weighting matrix

$$\begin{aligned}\mathbf{A} &= \mathbf{H}_{wall}^T \mathbf{Q}_{ibi} \mathbf{H}_{wall} \\ \mathbf{B} &= \mathbf{H}_{win}^T \mathbf{H}_{win},\end{aligned}\tag{2.15}$$

where \mathbf{Q}_{ibi} is a diagonal matrix

$$\mathbf{Q}_{ibi} = \text{diag} \left[\Delta, \Delta - 1, \dots, \mathbf{0}_{1 \times (v+1)}, 1, 2, \dots, n_c - v - \Delta \right]. \tag{2.16}$$

Another extension of the MSSNR method is the Minimum Delay Spread (MDS) algorithm [37]. This method minimizes the root mean square (RMS) delay spread of the effective channel, with a norm constraint on it. The \mathbf{A} and \mathbf{B} matrices are

$$\begin{aligned}\mathbf{A} &= \mathbf{H}_{wall}^T \mathbf{Q}_{mds} \mathbf{H}_{wall} \\ \mathbf{B} &= \mathbf{H}^T \mathbf{H},\end{aligned}\tag{2.17}$$

where \mathbf{Q}_{mds} is a diagonal matrix

$$\mathbf{Q}_{ibi} = \text{diag} \left[\eta^2, (\eta - 1)^2, \dots, 1, 0, 1, \dots, (n_c - \eta)^2 \right] \tag{2.18}$$

and where η is a design parameter indicating the desired *center of mass* or centroid of the channel impulse response.

2.1.3 Maximum Geometric Signal to Noise Ratio Method. In a multicarrier system, if the number of subchannels is large, it can be assumed that the subchannels are flat. Then each subchannel can be modeled as an Additive White Gaussian Noise (AWGN) channel. The achievable channel capacity of a multicarrier channel is the

sum of the capacities of the AWGN channels

$$b_{DMT} = \sum_{i \in \mathcal{S}} \log_2 \left(1 + \frac{SNR_i^{MFB}}{\Gamma} \right) \text{ bits/symbol}, \quad (2.19)$$

where i is the subchannel index, \mathcal{S} is the set of the indices of the used subchannels, SNR_i^{MFB} is the match filter bound of the i^{th} subchannel, and Γ is the SNR gap for achieving the Shannon channel capacity [38].

Al-Dhahir and Cioffi propose the Maximum Geometric Signal to Noise Ratio (MGSNR) method [27]. Their goal is maximizing the channel capacity based on the geometric SNR definition

$$\text{GSNR} = \Gamma \left(\left[\prod_{i \in \mathcal{S}} \left(1 + \frac{SNR_i^{EQ}}{\Gamma} \right) \right]^{1/N} - 1 \right). \quad (2.20)$$

Using (2.20), channel capacity can be rewritten as

$$b_{DMT} = N \log_2 \left(1 + \frac{\text{GSNR}}{\Gamma} \right) \text{ bits/symbol}, \quad (2.21)$$

which shows that all of the subchannels perform like N AWGN channels with each channel having SNR as GSNR. So, maximizing the GSNR means maximizing channel capacity. In (2.20) SNR_i^{MFB} is modified to include the effect of the equalizer.

Due to the use of several approximations, this algorithm does not truly maximize the bit rate. However, it is the first method that attempts to maximize the bit rate. Details of this method can be found in [27].

There are other equalizer design methods in the literature. Van Acker, *et al.* [33] propose a method called Per Tone Equalizer (PTE) using a bank of equalizer after the FFT to equalize a multicarrier system instead of a single filter. Similar to this method, Milosevic *et al.* [39] propose the Time Domain Equalizer Filter Bank (TEQ-FB).

Arslan *et al.* [23] propose a Maximum Bit Rate (MBR) design which very nearly maximizes the bit rate. However, like MGSNR, MBR also uses some approximations to maximize the bit rate, which make the algorithm not quite optimal.

2.2 Adaptive Channel Shorteners

Most of the channel shortener design methods in the literature assume the channel impulse response, including the methods in Section 2.1. This condition is seen in (2.6), (2.12) and (2.14) as observing the channel directly and can be managed by estimating the channel. Therefore, a training sequence should be sent with the transmitted signal to the receiver. Hence, the receiver estimates the channel with a copy of this training sequence and a decision criterion. But, sending the training sequence without information to estimate the channel reduces the bandwidth. For instance, 18% of the data in a time frame in a GSM system and 5% of an ADSL system is used for the training [40].

Also, if a channel changes faster, i.e. in wireless systems, training should be sent more frequently to estimate the channel effectively [41]. Thus, there are blind and semi-blind methods in channel equalization techniques. *Blind* means that a training sequence is not sent and the channel is estimated without using this training. Instead of a training sequence, statistics or structural characteristics of the transmitted signal are used in these systems. Since training is not used in the transmitted signals, more information can be sent in a frame. In semi-blind techniques, in order to track the frequent channel changes and to estimate the channel, relatively short training sequence or pilot tones are used. Semi-blind methods converge faster than (full) blind methods.

A decision about which signal characteristic or adaptive algorithm to use can be made in the design of blind or semi-blind methods. The most used structures in blind techniques for maximum likelihood estimation are moment, least mean square, maximum likelihood and minimum mean squared estimators. The most used adaptive algorithms are statistical gradient algorithms, recursive estimators, i.e., Recursive

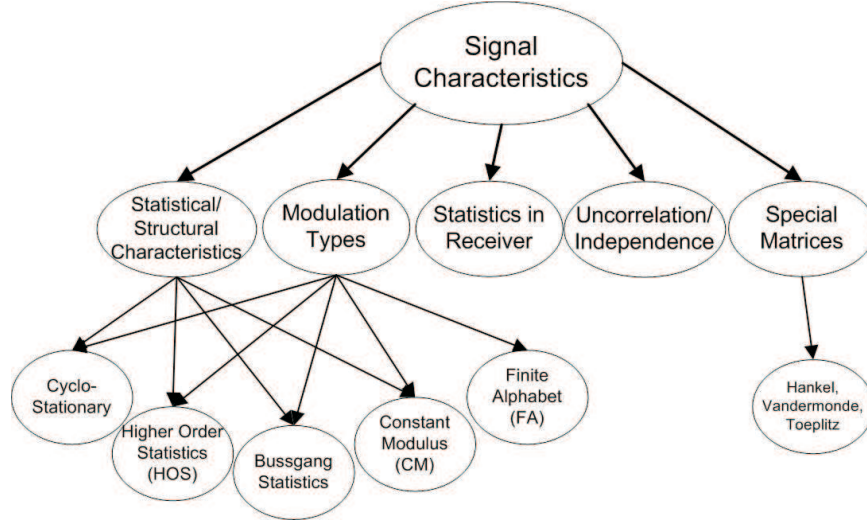


Figure 2.3: Blind signal processing methods [40].

Least Squares (RLS), Kalman filter, sub-space methods and expectation maximization algorithms. Some categories of signal characteristics used in blind techniques are shown in Fig 2.3.

While adaptive channel equalization is well-covered in the literature, adaptive channel shortening is not well-covered [41]. De Courville, *et al.* [42], propose a blind, adaptive equalizer for multicarrier receivers, but this design shortens the channel to a single spike instead of shortening the channel.

The most important channel shortening studies include the following. The Multicarrier Equalization by Restoration of Redundancy (MERRY) algorithm uses periodicity in a symbol due to the cyclic prefix [28]. The Sum-squared Autocorrelation Minimization (SAM) algorithm [29] minimizes the squared autocorrelation function of the shortened channel impulse response. The Single-Lag Autocorrelation Minimization (SLAM) algorithm [30] proposes a different solution in the context of the SAM and Blind Adaptive Channel Shortening Equalizer Algorithm which can Provide Shortened Channel State Information, (BACS-SI) [31] algorithm combines the advantages of the MMSE and SAM algorithms while eliminating their disadvantages.

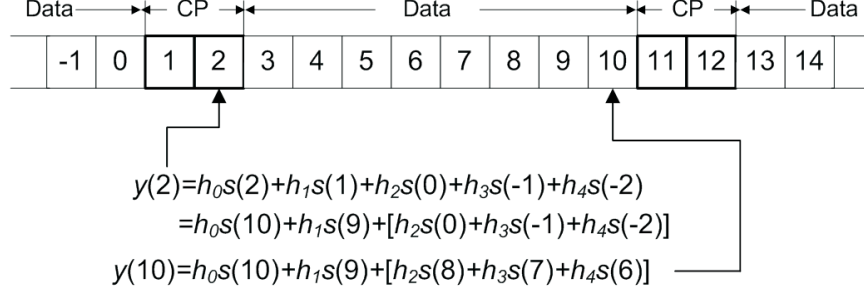


Figure 2.4: Difference in the ISI at the received CP and at the end of the received symbol [28].

2.2.1 Multicarrier Equalization by Restoration of Redundancy. The Multicarrier Equalization by Restoration of Redundancy (MERRY) algorithm uses redundancy which occurs due to the cyclic prefix. Thus, when v length cyclic prefix is added to the information signal which has N block length, the transmitted signal is periodic in $N + v$ samples. This periodicity is necessary to provide the orthogonality of subchannels [43]. After adding the CP, last and first v samples in a symbol are identical:

$$x(Mk + n) = x(Mk + n + N), \quad n \in \{1, 2, \dots, v\}, \quad (2.22)$$

where $M = N + v$ is the total symbol duration and k is the MCM symbol (block) index.

This idea can be seen in Fig. 2.4 for $N = 8$, $v = 2$ and $M = N + v = 10$ when $k = 0$. For channel impulse response \mathbf{h} , transmitted signal \mathbf{x} , noise \mathbf{n} equalizer \mathbf{w} and shortened channel, $\mathbf{c} = \mathbf{h} \star \mathbf{w}$, the received signal \mathbf{r} is

$$r(Mk + n) = \sum_{l=0}^{n_h} h_l x(Mk + n - l) + n(Mk + n), \quad (2.23)$$

and the output of the equalizer \mathbf{y} is

$$y(Mk + n) = \sum_{j=0}^{n_w} w_j r(Mk + n - j). \quad (2.24)$$

Because the different ISI effect to the CP and the last v samples of a symbol, the relationship in (2.22) is destroyed by ISI. For example, the 2nd and 10th samples are identical in Fig. 2.4. However, if the channel is longer than the CP, i.e. $n_h = 5$ and $v = 2$ in figure, the interfering samples before sample 2 are different from the interfering samples before sample 10. If h_2, h_3 and h_4 equal zero, then $y(2) = y(10)$. Hence, $y(2) = y(10)$ means $h_2 = h_3 = h_4 = 0$, and the effective channel is as short as CP. The MERRY cost function is

$$J_{MERRY} = E |y(Mk + v + \Delta) - y(Mk + v + N + \Delta)|^2, \quad \Delta \in \{0, \dots, M - 1\}, \quad (2.25)$$

where Δ represents delay. If a statistical gradient descent algorithm is applied to this cost function, the MERRY algorithm is:

Given Δ , for symbol $k = 0, 1, 2, \dots$

$$\begin{aligned} \tilde{\mathbf{r}}(k) &= \mathbf{r}(Mk + v + \Delta) - \mathbf{r}(Mk + v + N + \Delta) \\ e(k) &= \mathbf{w}^T(k) \tilde{\mathbf{r}}(k) \end{aligned} \quad (2.26)$$

$$\hat{\mathbf{w}}(k + 1) = \mathbf{w}(k) - \mu e(k) \tilde{\mathbf{r}}^*(k)$$

$$\mathbf{w}(k + 1) = \frac{\hat{\mathbf{w}}(k + 1)}{\|\hat{\mathbf{w}}(k + 1)\|},$$

where $\mathbf{r}(n) = \begin{bmatrix} r(n) & r(n - 1) & \dots & r(n - n_w) \end{bmatrix}^T$ and the $\|\mathbf{w}\| = 1$ constraint (unit norm constraint) is applied in the last step to eliminate the trivial solution $\mathbf{w} = \mathbf{0}$.

2.2.2 Sum-Squared Auto-correlation Minimization. Even though it is very simple, there are some disadvantages to MERRY algorithm, such the need for synchronization, the assumption a relatively time-invariant channel (since it updates symbol by symbol) and slow convergence. Alternatively, the Sum-Squared Auto-correlation Minimization (SAM) algorithm, which is based on suppressing the autocorrelation of the received signal outside a target window of size $v + 1$, is proposed [22, 29].

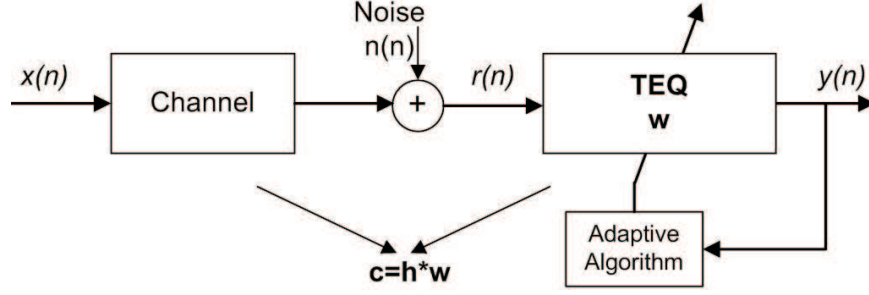


Figure 2.5: System model for SAM.

The system model is shown in Fig. 2.5, where $x(n)$, \mathbf{h} , $r(n)$, \mathbf{w} and $y(n)$ are shown as transmitted signal, FIR channel of length n_h , recieved signal, equalizer of length n_w and the equalized signal, respectively. The effective channel of length $n_c = n_h + n_w - 1$ is $\mathbf{c} = \mathbf{h} \star \mathbf{w}$. The TEQ is designed to shorten the effective channel \mathbf{c} to the length of CP plus one ($v + 1$).

If the received signal is

$$r(n) = \sum_{k=0}^{n_h} h(k)x(n-k) + n(n) \quad (2.27)$$

and the output of the TEQ is

$$y(n) = \sum_{k=0}^{n_w} w(k)r(n-k) = \mathbf{w}^T \mathbf{r}_n, \quad (2.28)$$

where $\mathbf{r}_n = \begin{bmatrix} r(n) & r(n-1) & \dots & r(n-n_w) \end{bmatrix}^T$, then the autocorrelation function is

$$R_{cc}(l) = \sum_{k=0}^{n_c-1} c(k)c(k-l). \quad (2.29)$$

Being zero for the outside coefficients of a window of length $v + 1$ of the effective channel \mathbf{c} , and being zero for the outside autocorrelation values of length $2v + 1$ of R_{cc} , are identical:

$$R_{cc}(l) = 0, \quad \forall |l| > v. \quad (2.30)$$

The constraints, $\|\mathbf{c}\|_2^2 = 1$ or $R_{cc}(0) = 1$, are applied to the effective channel or the autocorrelation function to avoid the trivial solution. The SAM cost function is defined as

$$J_{SAM} = \sum_{l=v+1}^{n_c-1} |R_{cc}(l)|^2. \quad (2.31)$$

If we assume the noiseless scenario, the cost function is

$$J_{SAM} = \sum_{l=v+1}^{n_c-1} |R_{cc}(l)|^2 = \sum_{l=v+1}^{n_c-1} |R_{yy}(l)|^2, \quad (2.32)$$

The steepest descent algorithm according to (2.32) is

$$\mathbf{w}^{new} = \mathbf{w}^{old} - \mu \nabla_{\mathbf{w}} \left(\sum_{l=v+1}^{n_c-1} |R_{yy}(l)|^2 \right), \quad (2.33)$$

where μ is step size and ∇ denotes gradient with respect to \mathbf{w} . The expectation operation can be replaced with Moving Average (MA) or Auto-regressive (AR) estimates to implement the algorithm. This impementation and the details of the SAM algorithm are in [22, 29].

2.2.3 Single Lag Auto-correlation Minimization. For the SAM algorithm, suppressing the outside coefficients of a target window of the autocorrelation function means keeping the effective channel length in this window. Thus, the first outside coefficient of a target window of autocorrelation function gives the effective channel of that length. In other words, looking at the first coefficient outside a target window shows whether the channel is shortened or not. The Single Lag Auto-correlation Minimization (SLAM) algorithm uses this idea [30].

The system model is the same as Fig. 2.5 and the cost function is

$$J_{SLAM} = |R_{cc}(l)|^2, \quad l = v + 1, \quad (2.34)$$

where just $l = v + 1$ lag of the autocorrelation function is minimized.

SLAM is a variation of SAM and keeps all the advantages and disadvantages of SAM, but has lower complexity.

2.2.4 Blind Adaptive Channel Shortening Equalizer Algorithm which can Provide Shortened Channel State Information. Due to a major shortcoming of the blind algorithms (MERRY, SAM, SLAM), Toker and Altın proposed the Blind Adaptive Channel Shortening Equalizer Algorithm which can provide Shortened Channel State Information (BACS-SI) [31] to provide a shortened channel impulse response after shortening.

Although the blind algorithms mentioned above develop TEQ and successfully shorten the channel, they do not provide the shortened channel impulse response after shortening, which is *necessary* for a proper design of an FEQ. In contrast, the proposed BACS-SI algorithm can both shorten the channel in a blind manner and also explicitly provides the shortened channel impulse response.

As a result of the blind nature of the problem, the optimal solution (e.g., in the sense of maximum bit rate) cannot be directly found due to the lack of prior channel knowledge. Hence, the algorithm is proposed in two phases. In the first phase, a stochastic gradient descent algorithm is derived to find an arbitrary minimum of the surface. In the second phase of the algorithm, a global search technique is proposed based on Genetic Algorithm (GA).

This algorithm combines the advantages of the MMSE and the SAM algorithm, while eliminating their disadvantages. The system model is shown in Fig. 2.1.

Similar to analysis in (2.32), the autocorrelation of the sequence $\hat{y}(n)$ is (under the i.i.d. assumption)

$$R_{\hat{y}\hat{y}}(l) = \sum_{k=0}^{n_b-1} b(k)b(k-l) = R_{bb}(l). \quad (2.35)$$

Using the definitions in (2.32) and (2.35) the cost function of the algorithm is (again in noiseless senario)

$$J_{BACS} = \sum_{l=0}^{n_c-1} (R_{yy}(l) - R_{\widehat{yy}}(l))^2 = \sum_{l=0}^{n_c-1} (R_{cc}(l) - R_{bb}(l))^2 \quad (2.36)$$

In order to avoid trivial solution, the unit norm constraint on the equalizer, $\mathbf{w}^T \mathbf{w} = 1$, is used. To find the minimums of the cost surface, a stochastic gradient descent algorithm is applied:

$$\begin{aligned} \mathbf{f}^{n+1} &= \mathbf{f}^n - \frac{1}{2} \mu \nabla_{\mathbf{f}} J(n) \\ \nabla_{\mathbf{f}} J(n) &= \begin{bmatrix} \nabla_{\mathbf{w}} J(n) \\ \nabla_{\mathbf{b}} J(n) \end{bmatrix}, \end{aligned} \quad (2.37)$$

where μ is the step size and $\nabla_{\mathbf{f}}$, $\nabla_{\mathbf{w}}$ and $\nabla_{\mathbf{b}}$ are gradients of the cost function with respect to \mathbf{f} , \mathbf{w} and \mathbf{b} , respectively. These gradients of the $J_{BACS-SI}$ w.r.t. \mathbf{w} and \mathbf{b} are

$$\nabla_{\mathbf{w}} J(n) = 2 \sum_{l=0}^{L_c} (R_{yy}(l) - R_{bb}(l)) \nabla_{\mathbf{w}} R_{yy}(l) \quad (2.38)$$

and

$$\nabla_{\mathbf{b}} J(n) = 2 \sum_{l=0}^{L_c} (R_{yy}(l) - R_{bb}(l)) (\mathbf{b}_{up}(l) + \mathbf{b}_{dn}(l)), \quad (2.39)$$

where the $\mathbf{b}_{up}(l)$ and $\mathbf{b}_{dn}(l)$ vectors are defined as $\mathbf{b}_{up}(l) = [b_l \ \cdots \ b_{n_b-1} \ \mathbf{0}_{1 \times l}]^T$ and $\mathbf{b}_{dn}(l) = [\mathbf{0}_{1 \times l} \ b_0 \ \cdots \ b_{n_b-l-1}]^T$.¹

Because the cost surface is multimodal with groups of minima (related to each other) having identical costs, a stochastic gradient descent algorithm is derived to find one of these minima in order to access the others. Although the minima optimize the BACS-SI cost surface, only one of them actually performs the best shortening (in the sense of criterion (2.40) below). In the second phase of the algorithm, a GA based

¹For the expectation terms, MA and AR estimates are the same as in the SAM algorithm. Details of the algorithm and $\mathbf{b}_{up}(l)$ and $\mathbf{b}_{dn}(l)$ derivations are in [31]

method is proposed to find this minimum and the corresponding equalizer and TIR taps, considering a fitness function derived from the pilot tone utilized in ADSL:

$$\min_{(\mathbf{w}, \mathbf{b})} \sum_{k \in K} |R(k) - 1|^2, \quad (2.40)$$

where K is the set of pilot tones define in [44]. Actually, the 64th tone in the ADSL standard is reserved as a pilot signal for timing recovery and is set to 1, i.e. $R(64) = 1$ [44, 45].

2.3 Genetic Algorithm

The Genetic Algorithm (GA) is an optimization method based on the principles of natural selection and genetics. GA is first developed by John Holland [46] and was improved by his student David Goldberg [47]. A GA evolves a population composed of many individuals under some selection rules to maximize the *fitness* (or minimize the cost). In every iteration, the strongest/fittest individual is kept inside the population and the optimum solution is found by generating new children from these best parents. Because the new generation has some specifications from their parents, the next step is better and the solution can be reached quickly.

Some of the advantages of a GA include that it

- Optimizes with continuous or discrete variables,
- Doesn't require derivative information,
- Simultaneously searches from a wide sampling of the cost surface,
- Deals with a large number of variables,
- Is well studied for parallel computers,
- Optimizes variables with extremely complex cost surfaces (they can jump out of a local minimum),
- Provides a list of optimum variables, not just a single solution,
- May encode the variables so that the optimization is done with the encoded variables, and
- Works with numerically generated data, experimental data or analytical functions [48].

The GA is not the best method to solve every problem. For instance, for a convex analytical function of only a few variables, other methods are relatively better and find the solution quickly while the GA is still analyzing the costs of initial population. However, the advantages listed above (especially avoiding local minima) recommend the GA for higher level problem solutions and global search techniques [48].

Defining the optimization variables, the cost function and the cost are the first steps in starting a GA. The GA ends by testing convergence, as in other optimization techniques. A flowchart of a GA is shown in Fig. 2.6.

2.3.1 Selecting the Variables and the Cost Function. A cost function may depend on just one variable, such as a mathematical function or an experiment, or it may depend on many variables. After defining the cost function, the variables to be optimized are converted to a continuous or binary system (for a binary GA or continuous GA) called *genes*. In binary GA, the variables (genes) are coded to 0 or 1. *Choromosomes* can be formed by combining the genes. A chorosome which is formed of N_{var} genes of N_{gene} length is shown below

$$chorosome = \left[\underbrace{1111001001}_{gene_1} \underbrace{0011011111}_{gene_2} \dots \underbrace{0000101001}_{gene_{N_{var}}} \right].$$

2.3.2 Initial Population. The GA starts with a group of chorosomes created with random numbers known as the initial population. Usually population size is determined by users. If the population size is large enough, a detailed search of the cost surface can be managed. Each row in the population matrix corresponds to a chromosomes, and these chromosome are sent to the cost function for evaluation.

2.3.3 Natural Selection. Survival of the fittest in nature means to keep the chorosomes with the lowest cost/highest fitness in the population in GA. Before selection, chromosomes are sorted lowest to highest (or highest to lowest if the search

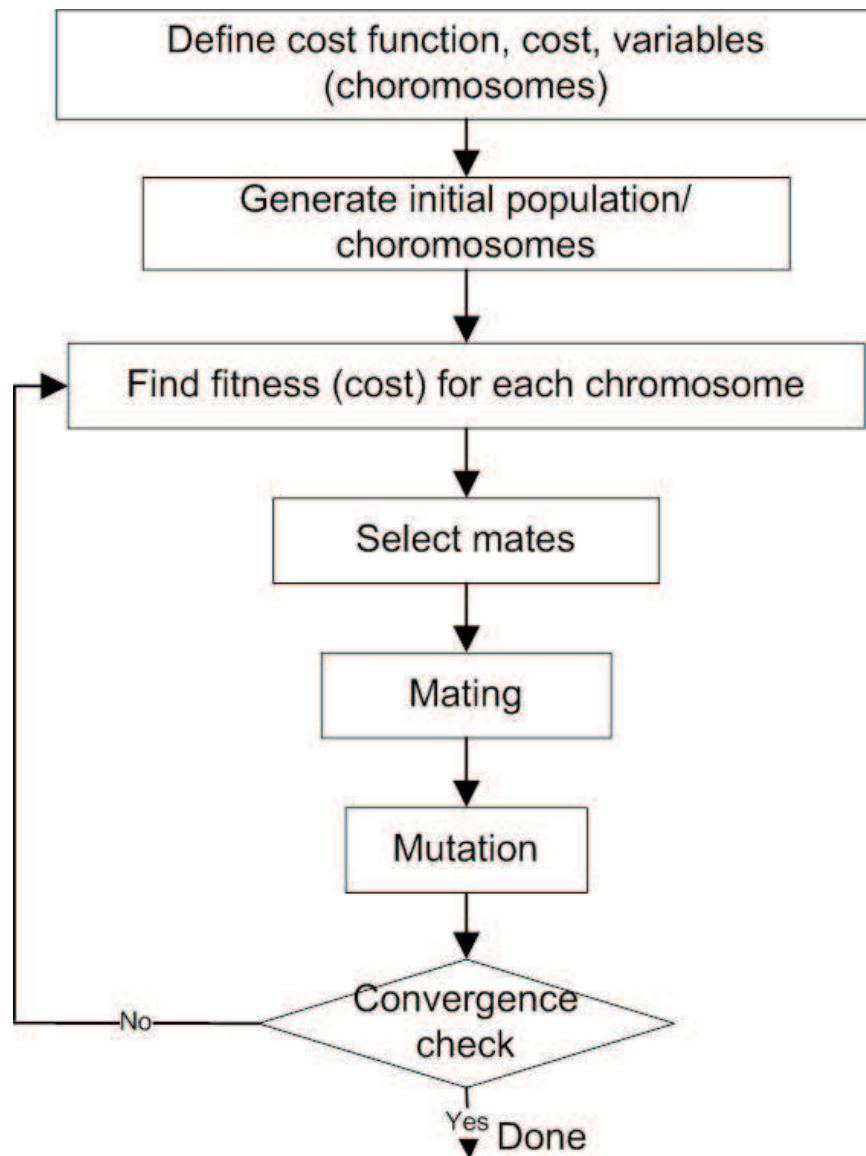


Figure 2.6: Flowchart of a GA.

is for a maximum)². The user defined selection rate determines the survivor number (only the best are selected) for the mating. Later a pair is selected to generate new individuals from the chromosomes that are kept. The pair selection methods can be different [48]:

1. Selecting the best and the worst (top and bottom) based on their costs
2. Random selection
3. Weighted random selection. The selection probability of the best chromosome is higher.
4. Tournament selection. Pairs are selected from two or more subsets.

2.3.4 Mating. Mating is one of the GA steps that creates new children from the parents. The most common form of the mating process randomly selects a crossover point between the first and the last digit of parent chromosomes and switches the remaining parts of the chromosomes with each other. As seen in Fig. 2.7, the first part of the *parent*₁ and the second part of the *parent*₂ generate the *child*₁ and the second part of the *parent*₁ and the first part of the *parent*₂ generate the *child*₂. Mating can be achieved by different methods. For example, two or more crossover points can be selected, and parts of the chromosomes between crossover points can be switched among them.

2.3.5 Mutation. The sequence of a gene can be altered in many ways, including environmental factors, living standards etc. Mutation keeps the algorithm from converging too fast to local minima before searching the whole surface. Mutation alters the randomly selected bits of randomly selected chromosomes from 1 to 0 or 0 to 1 according to a user-specified mutation rate. Also in this step, selection can be applied randomly or can be user selected. But, applying mutation to many bits increases convergence time.

²After this point, only minimizing a function based on the binary GA will be considered.

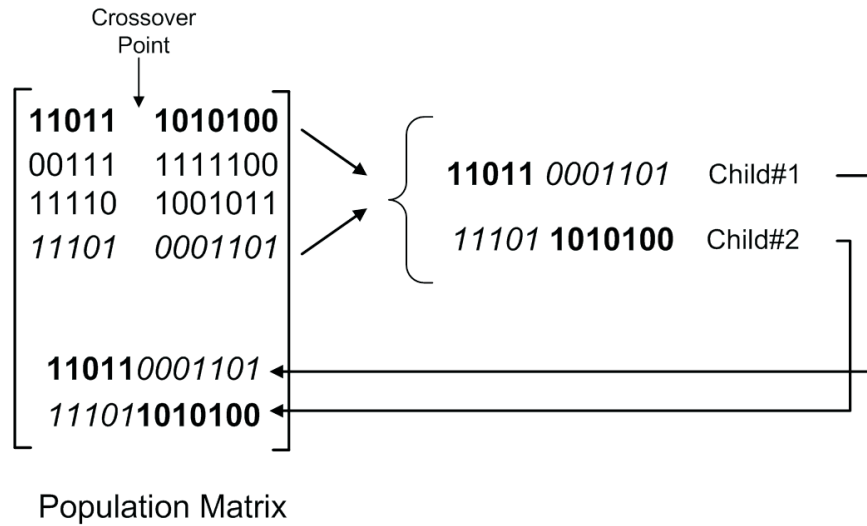


Figure 2.7: Two parents mate to generate two children. Then, these two children are added to the population.

2.3.6 The Next Generation and Convergence. After mutation, the algorithm goes to the first step to produce a new population and better individuals. After some iterations, chromosomes remain the same and the cost function does not change. At this point, the algorithm should be stopped. The algorithm is stopped when the predetermined iteration number or the minimum cost is reached. Every other iteration, the minimum cost or iteration number is checked for stopping the algorithm.

III. Bit-Error-Rate-Minimizing Channel Shortening using Post-FEQ Diversity Combining and a Genetic Algorithm

In cyclic prefixed systems, the channel delay spread should be shorter than the CP. If so, the equalization can be done by a frequency domain equalizer (FEQ), which is just a bank of complex scalars. If not, a Channel Shortening Equalizer (CSE) or Time domain Equalizer (TEQ) is used at the receiver front-end to shorten the channel to the desired length. Once a shortened channel is obtained, equalization can be done straightforwardly.

DMT systems, especially digital subscriber lines, initially gained prominence in wireline applications. In DSL, system performance is measured by bit rate not exceeding a given bit error rate (BER). Thus, previous works have attempted to maximize the bit rate [17, 23, 26, 27, 32]. In contrast, OFDM systems have a fixed bit loading algorithm at the transmitter, and the performance metric is minimization of the BER.

Since the channel changes more frequently in wireless systems, the adaptive channel shorteners and FEQs are more important. Adaptive channel shortening has been examined in many papers [28–31]. However, the adaptive FEQ design has not been examined except in [49]. Also, most channel shortening and frequency domain equalizer algorithms assume knowledge of the channel via a training signal, but in time-varying environments, the training data must be sent over again, which reduces the throughput of the system.

This chapter has two sections. In section 3.1, the system architecture is modified to obtain a lower BER in the context of [49]. The recursive least squares (RLS)-like implementation of the Multicarrier Equalization by Restoration of Redundancy (MERRY) channel shortener is used to shorten the channel, and three types of adaptation rules, i.e., DD-Least Mean Squared (LMS), DD-Recursive Least Squares (RLS), and Constant Modulus Algorithm (CMA), are used to adapt the FEQs. To provide the minimum BER, different types of diversity combining are considered.

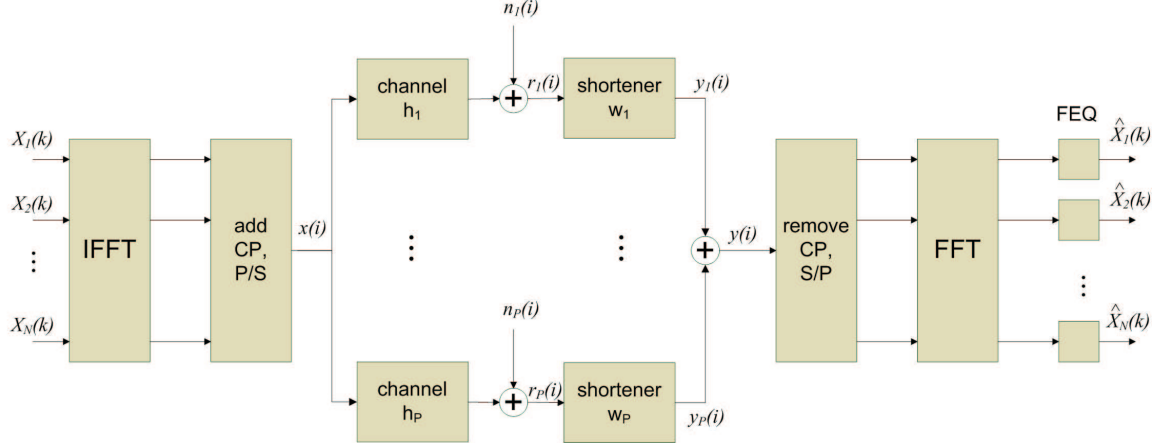


Figure 3.1: SIMO multicarrier system model.

In section 3.2, a CSE design is used that attempts to minimize the BER at the output of the traditional receiver using Genetic Algorithm (GA), which is a global search method based on the principles of natural selection and genetics.

3.1 BER Minimizing Channel Shortening using Post-FEQ Diversity Combining

First assume a Single-Input Multiple-Output (SIMO) multicarrier system with different post-FEQ diversity combining techniques. Figs. 3.1, 3.2 and 3.3 show a typical SIMO multicarrier system and the proposed post-FEQ diversity combining techniques. Fig. 3.1 shows the traditional architecture and has 1 FFT, and Figs. 3.2 & 3.3 show proposed variations that have 2 FFTs and P FFTs. All 3 have P receive antennas. Traditional receiver architecture cannot restore the values of the tones affected by deep fades. On the other hand, new architectures may combine different paths to obtain the values of tones.

The proposed architectures generalize the existing architecture by moving the diversity combining point from before the FFT/FEQ to after the FFT/FEQ, thus allowing for different FEQ coefficients for each path.

The frequency domain bins (generally modulated with quadrature amplitude-modulation (QAM)) are blocked into groups of size N . In order to obtain the trans-

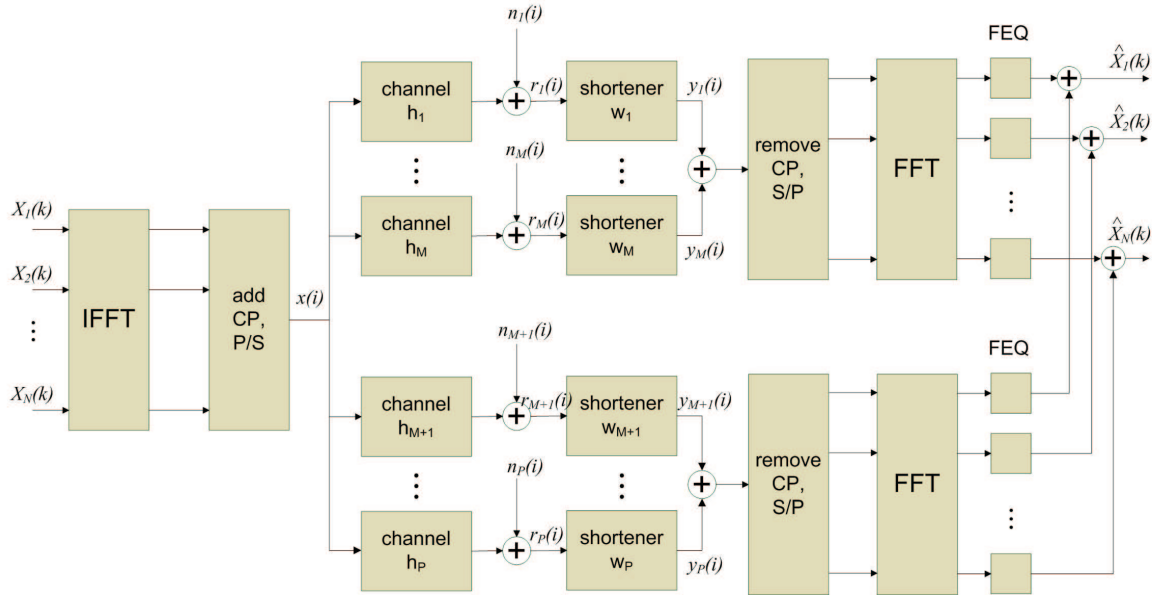


Figure 3.2: Post-FEQ diversity combining with 2 FFT.

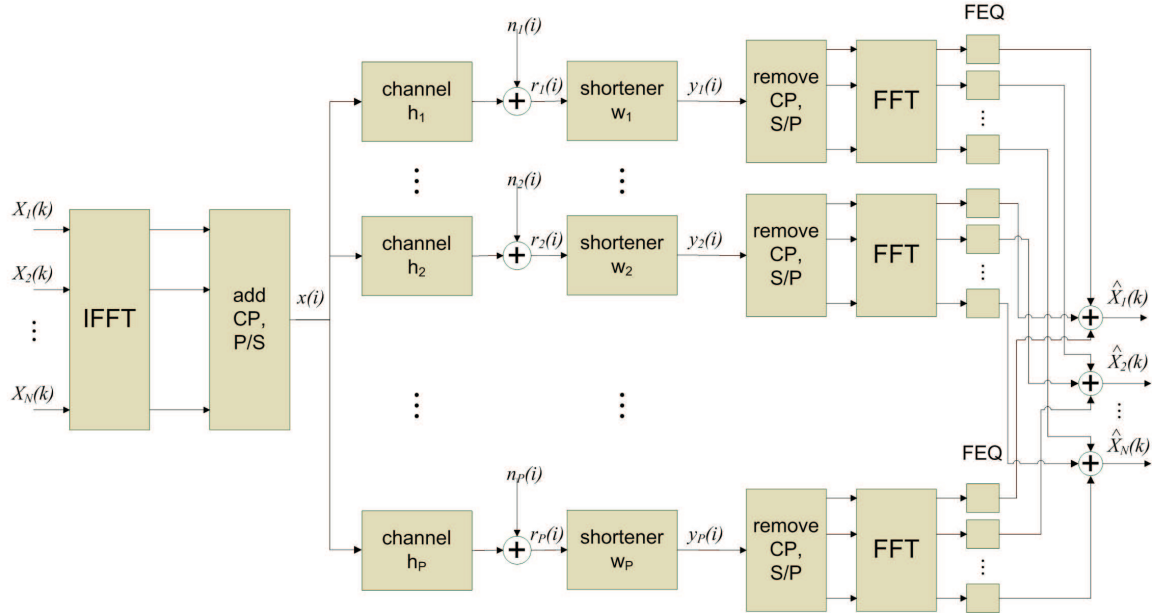


Figure 3.3: Post-FEQ diversity combining with individual FFT.

mitted signal, these blocks, e.g., the k^{th} block $\mathbf{X}(k)$, are transformed into the time domain by taking an inverse fast Fourier transform (IFFT). The last v samples of each block are appended to the beginning of the block as a cyclic prefix. Then these $M = N + v$ samples are transmitted serially. Since the CP is added, the first v transmitted data samples, denoted $x(i)$, equal the last v samples, i.e.,

$$x(Mk + i) = x(Mk + i + N), \quad i \in 1, 2, \dots, v. \quad (3.1)$$

The received data from the p^{th} path r_p is

$$r_p(i) = \sum_{l=0}^{L_h-1} h_p(l)x(i-l) + n_p(i), \quad p \in 1, \dots, P, \quad (3.2)$$

and the equalized data, i.e., the output of the p^{th} CSE, is

$$\begin{aligned} y_p(i) &= \sum_{l=0}^{L_w-1} w_p(l)r_p(i-l) \\ &= \sum_{l=0}^{L_c-1} c_p(l)x(i-l) + \sum_{l=0}^{L_w-1} w_p(l)n_p(i-l). \end{aligned} \quad (3.3)$$

This output is shown as the final output of the system in Fig. 3.3. For the systems in Fig. 3.1 and Fig. 3.2, outputs can be obtained straightforwardly as

$$y(i) = \sum_{p=1}^P y_p(i) \quad (3.4)$$

and

$$y^1(i) = \sum_{p=1}^{P/2} y_p(i) \quad (3.5a)$$

$$y^2(i) = \sum_{p=P/2+1}^P y_p(i). \quad (3.5b)$$

The channel h_p , the CSE w_p , and the shortened channel $c_p = h_p \star w_p$, have lengths L_h , L_w , and $L_c = L_h + L_w - 1$.

3.1.1 Review of the RLS-MERRY Algorithm. In this section, the RLS-MERRY algorithm is reviewed, and extended to these new architectures. The original implementation of MERRY algorithm (described in Chapter 2) [28] is an LMS-like algorithm with a unit norm constraint. In [49], the MERRY algorithm is reformulated from an eigenvector problem with a unit norm constraint, (UNC), into a least-squares problem with a unit tap constraint, (UTC), to enable an RLS algorithm instead of LMS. Additionally, in order to implement a “*blind*” RLS algorithm to minimize the MERRY cost function, a pseudo-*desired* signal is introduced. As indicated in Chapter 2, the cost function of MERRY is

$$J_{MERRY} = E [|y(Mk + v + \Delta) - y(Mk + v + N + \Delta)|^2]$$

$$\Delta \in \{0, \dots, M - 1\}, \quad (3.6)$$

where Δ is the delay.

The error is defined as

$$y(Mk + v + \Delta) - y(Mk + v + N + \Delta) = \mathbf{w}^T \tilde{\mathbf{r}}(k), \quad (3.7)$$

where

$$\mathbf{w}^T = [\mathbf{w}_1^T, \mathbf{w}_2^T, \dots, \mathbf{w}_P^T], \quad (3.8)$$

$$\tilde{\mathbf{r}}^T(k) = [\tilde{\mathbf{r}}_1^T(k), \tilde{\mathbf{r}}_2^T(k), \dots, \tilde{\mathbf{r}}_P^T(k)] \quad (3.9)$$

which consists of

$$\tilde{\mathbf{r}}_p(k) = \begin{bmatrix} r_p(Mk + v + \Delta) \\ \vdots \\ r_p(Mk + v + \Delta - L_w) \end{bmatrix} - \begin{bmatrix} r_p(Mk + v + N + \Delta) \\ \vdots \\ r_p(Mk + v + N + \Delta - L_w) \end{bmatrix}$$

for $1 \leq p \leq P$.

The RLS-like algorithm does not have a *desired* or *training* signal for comparison. However, using UTC instead of UNC addresses this problem. The vector \mathbf{w} is truncated by removing $w(i_p)$, where $i_p \in [0, L_w - 1]$:

$$\mathbf{w}_t = [w_p(0), w_p(1), \dots, w_p(i_p - 1), w_p(i_p + 1), \dots, w_p(PL_w - 1)]^T, \quad (3.10)$$

and similarly for $\tilde{\mathbf{r}}_t(k)$. To simplify the notation, consider the P CSEs, and the CSE that has a tap constrained to unity to be $p = 1$, so that $w(i_1) = 1$. Then the cost function with UTC is

$$J_{MERRY} = E [\mathbf{w}^T \tilde{\mathbf{r}}(k) \tilde{\mathbf{r}}^H(k) \mathbf{w}^*] \quad (3.11)$$

$$= E \left[\left| \mathbf{w}_t^T \tilde{\mathbf{r}}_t(k) + 1 \cdot \tilde{\mathbf{r}}_{i_1}(k) \right|^2 \right] \quad (3.12)$$

$$\triangleq E \left[\left| \mathbf{w}_t^T \tilde{\mathbf{r}}_t(k) - d(k) \right|^2 \right], \quad (3.13)$$

where the desired signal is

$$d(k) = -\tilde{\mathbf{r}}_{i_1}(k) \quad (3.14)$$

$$= r_1(Mk + v + N + \Delta - i_1) - r_1(Mk + v + \Delta - i_1).$$

Finally, the RLS algorithm is

$$e(k) = (r_1(Mk + v + N + \Delta - i_1) \quad (3.15)$$

$$- r_1(Mk + v + \Delta - i_1)) - \mathbf{w}_t^T(k) \tilde{\mathbf{r}}_t(k)$$

$$\mathbf{z}(k) = \mathbf{R}^{-1}(k) \tilde{\mathbf{r}}_t^*(k) \quad (3.16)$$

$$\tilde{\mathbf{z}}(k) = \frac{\mathbf{z}(k)}{\rho + \tilde{\mathbf{r}}_t^T(k) \mathbf{z}(k)} \quad (3.17)$$

$$\mathbf{w}_t(k+1) = \mathbf{w}_t(k) + e(k) \tilde{\mathbf{z}}(k) \quad (3.18)$$

$$\mathbf{R}^{-1}(k+1) = \frac{1}{\rho} (\mathbf{R}^{-1}(k) - \tilde{\mathbf{z}}(k) \mathbf{z}^H(k)). \quad (3.19)$$

Here ρ stands for the “forgetting factor,” where all data has the same effect when $\rho = 1$, and the recent data has further effect when ρ is smaller. The \mathbf{R}^{-1} is initialized as an identity matrix, i.e., $\eta \mathbf{I}_{(PL_w-1)}$, where η is a large positive constant.

The RLS-MERRY algorithm for Fig. 3.1 is described above. In Figs. 3.2 & 3.3, the algorithm changes slightly, such as in applying the UTC. For the 2 FFT system in Fig. 3.2, the CSEs that have a UTC are $p = 1$ and $p = P/2 + 1$, and the i_1^{th} and the i_2^{th} elements are 1, so that $0 \leq i_1 \leq L_w - 1$, $P/2L_w \leq i_2 \leq (P/2 + 1)L_w - 1$.

For the P FFT system in Fig. 3.3, each of the P CSEs has a UTC, so that $0 \leq i_1 \leq L_w - 1$, $L_w \leq i_2 \leq 2L_w - 1$, etc. The truncated vectors \mathbf{w}_t (and similarly for $\tilde{\mathbf{r}}_t(k)$), for the 2 FFT and P FFT systems are

$$\mathbf{w}_t^1 = [w_1(0), \dots, w_1(i_1 - 1), w_1(i_1 + 1), \dots, w_1(P/2L_w - 1)] \quad (3.20)$$

$$\mathbf{w}_t^2 = [w_{P/2+1}(P/2L_w), \dots, w_{P/2+1}(i_2 - 1), w_{P/2+1}(i_2 + 1), \dots, w_{P/2+1}(PL_w - 1)]^T \quad (3.21)$$

$$\mathbf{w}_t^1 = [w_1(0), \dots, w_1(i_1 - 1), w_1(i_1 + 1), \dots, w_1(L_w - 1)] \quad (3.22a)$$

$$\mathbf{w}_t^2 = [w_2(L_w), \dots, w_2(i_2 - 1), w_2(i_2 + 1), \dots, w_2(2L_w - 1)] \quad (3.22b)$$

\vdots

$$\mathbf{w}_t^P = [w_P((P - 1)L_w), \dots, w_P(i_p - 1), w_P(i_p + 1), \dots, w_P(PL_w - 1)]^T. \quad (3.22c)$$

The desired signal for (3.13) is

$$\begin{aligned} d_{\#FFT}(k) &= -\tilde{\mathbf{r}}_{i_{\#FFT}}(k) \\ &= r_p(Mk + v + N + \Delta - i_{\#FFT}) - r_p(Mk + v + \Delta - i_{\#FFT}). \end{aligned} \quad (3.23)$$

Note that there are two desired signals for the 2 FFT system and P desired signals for the P FFT system. The error term in (3.15) and the update rule in (3.18) are re-written based on these truncated and desired vectors for the RLS algorithm of the 2 FFT and P FFT systems.

3.1.2 Adaptive FEQ. After shortening the channel, an FEQ is used to eliminate complex-valued flat fading on each tone at the receiver. Since FEQ has an important role in equalization, it can be updated with an adaptive algorithm. Designing a system which gives a minimum BER, while updating the FEQ is the primary objective of this work. There are two different new post-FEQ combining techniques, as shown in Fig. 3.2 and Fig. 3.3, that enable a lower BER. The proposed new architectures have an opportunity to combine the values of tones after the IFFT which are affected by deep fades. For example, assume two different channel responses as in Fig. 3.4. A tone value can be affected by a response as in Fig. 3.4(a). Also, the same tone value can be affected by a deep fade as in Fig. 3.4(b). Thus, diversity combining provides independent samples of each tone. Finally, note that more diversity as in Fig. 3.3 provides robustness to fades and lower BER, since it is unlikely that both samples of a tone will contain a fade. Thus, the FEQ adaptation algorithms, i.e., LMS, RLS and CMA, are extended to the new diversity combining architectures.

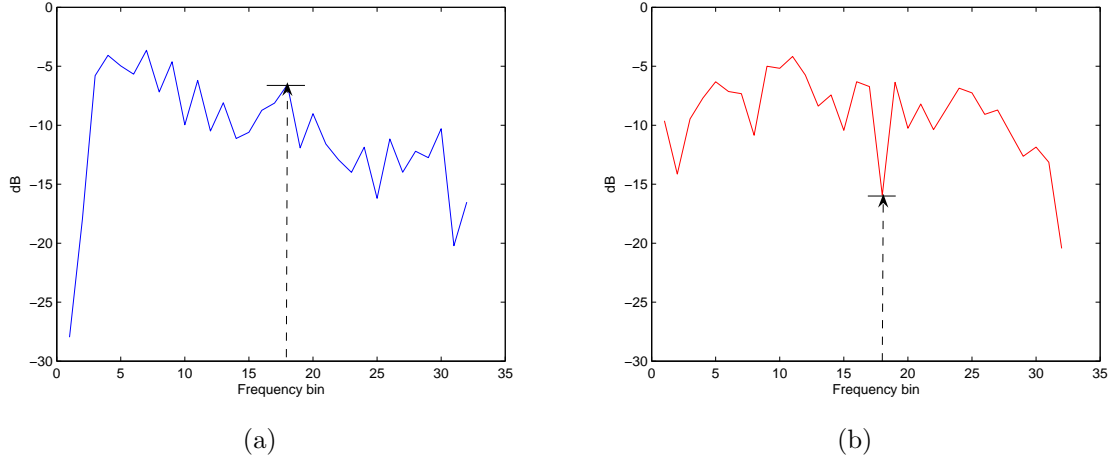


Figure 3.4: A tone can be affected by different types of fades as in the channels (a) and (b).

3.1.2.1 Least Mean Square Algorithm. The Least Mean Square (LMS) algorithm is an adaptive algorithm which uses a gradient-based *method of steepest descent*. LMS makes corrections to the weight vector in the direction of the gradient vector, which eventually leads to the minimum mean square error. The LMS algorithm is relatively simple compared to other algorithms, i.e., it does not require calculation of a correlation function or matrix inversion [16].

From the steepest descent algorithm, the adaptation rule is

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{1}{2}\mu[-\nabla(E\{e^2(n)\})] \quad (3.24)$$

where μ is the step size, and $e(n)$ is the error between the output of the system and the reference (desired) signal, which is

$$e(n) = [d^*(n) - \mathbf{w}^H \mathbf{x}(n)]. \quad (3.25)$$

The gradient vector is

$$\nabla_{\mathbf{w}}(E\{e^2(n)\}) = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n), \quad (3.26)$$

where \mathbf{p} is the cross-correlation vector and \mathbf{R} is the correlation matrix. In the steepest descent algorithm, \mathbf{p} and \mathbf{R} must be obtained in real time. The LMS algorithm simplifies this problem by using instantaneous values instead of actual values of \mathbf{p} and \mathbf{R} , i.e.,

$$\mathbf{R} = \mathbf{x}(n)\mathbf{x}^H(n) \quad (3.27)$$

$$\mathbf{p} = d^*(n)\mathbf{x}(n). \quad (3.28)$$

Finally, the weight update rule is

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\mathbf{w}(n)] \quad (3.29)$$

$$= \mathbf{w}(n) + \mu\mathbf{x}(n)e^*(n). \quad (3.30)$$

The LMS algorithm is initialized as $w(0)$ for the weight vector at $n = 0$. Thus, the LMS algorithm is summarized as follows [16]:

$$\text{Output,} \quad y(n) = \mathbf{w}^H(n)\mathbf{x}(n) \quad (3.31)$$

$$\text{Error,} \quad e(n) = d^*(n) - y(n) \quad (3.32)$$

$$\text{Weight,} \quad \mathbf{w}(n+1) = \mathbf{w}(n) + \mu\mathbf{x}(n)e^*(n). \quad (3.33)$$

3.1.2.2 Recursive Least Squares Algorithm.

The Recursive Least Squares (RLS) algorithm finds adaptive filter coefficients that relate to recursively producing a least squares metric of the error signal [16]. This algorithm depends on the signals; the minimum mean square error (MMSE) algorithm, however, depends on statistics. An important feature of the RLS algorithm is its convergence rate. The convergence rate of RLS is an order of magnitude faster than the LMS algorithm because the RLS algorithm *whitens* the input signal by using the matrix inversion.

The weighted least squares error function is

$$\sum_{i=1}^n \rho^{n-i} |e(i)|^2 = \sum_{i=1}^n \rho^{n-i} |d(i) - \mathbf{w}^H(n)\mathbf{x}(i)|^2 \quad (3.34)$$

where $0 < \rho \leq 1$ is an exponential weighting factor (forgetting factor) and the error e is same as in (3.25). This function is minimized by taking the gradient with respect to \mathbf{w} and setting it to zero:

$$\nabla_{\mathbf{w}} \left(\sum_{i=1}^n \rho^{n-i} |e(i)|^2 \right) = 0. \quad (3.35)$$

The cross-correlation vector $\mathbf{z}(n)$ is then

$$\mathbf{z}(n) = \sum_{l=1}^p w_n(l) \left[\sum_{i=1}^n \rho^{n-i} x(i-l)x^*(i-k) \right] = \sum_{i=1}^n \rho^{n-i} d(i)x^*(i-k) \quad (3.36)$$

This form is also

$$\mathbf{R}(n)\mathbf{w}(n) = \mathbf{z}(n), \quad (3.37)$$

where $\mathbf{R}(n)$ is the weighted auto-correlation matrix. The filter taps are then

$$\mathbf{w}(n) = \mathbf{R}^{-1}(n)\mathbf{z}(n). \quad (3.38)$$

In order to compute the least-square estimate for $\mathbf{w}(n)$, the inverse of the correlation matrix must be determined. This determination uses the *matrix inversion lemma* [16]. Finally, the inverse correlation matrix is

$$\begin{aligned} \mathbf{P}(n) &= \mathbf{R}^{-1}(n) \\ &= \rho^{-1}\mathbf{P}(n-1) - \mathbf{g}(n)\mathbf{x}^T(n)\rho^{-1}\mathbf{P}(n-1), \end{aligned} \quad (3.39)$$

where the *gain vector* $\mathbf{g}(n)$ is

$$\mathbf{g}(n) = \rho^{-1} \mathbf{P}(n-1) \mathbf{x}^*(n) \{1 + \mathbf{x}^T(n) \rho^{-1} \mathbf{P}(n-1) \mathbf{x}^*(n)\}^{-1} \quad (3.40)$$

$$= \mathbf{P}(n-1) \mathbf{x}^*(n) \{\rho + \mathbf{x}^T(n) \mathbf{P}(n-1) \mathbf{x}^*(n)\}^{-1}. \quad (3.41)$$

This vector is

$$\mathbf{g}(n) = \mathbf{P}(n) \mathbf{x}^*(n). \quad (3.42)$$

Finally, the recursive solution of the weighted tap vector is

$$\begin{aligned} \mathbf{w}(n) &= \mathbf{P}(n) \mathbf{z}(n) \\ &= \rho \mathbf{P}(n) \mathbf{z}(n-1) + d(n) \mathbf{P}(n) \mathbf{x}^*(n) \end{aligned} \quad (3.43)$$

with

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{g}(n) [d(n) - \mathbf{x}^T(n) \mathbf{w}(n-1)] \quad (3.44)$$

$$= \mathbf{w}(n-1) + \mathbf{g}(n) \alpha(n), \quad (3.45)$$

where $\alpha(n) = d(n) - \mathbf{x}^T(n) \mathbf{w}(n-1)$ is the *a priori* error. Comparing with the *a posteriori* error, the error calculated after the filter is updated:

$$e(n) = d(n) - \mathbf{x}^T(n) \mathbf{w}(n). \quad (3.46)$$

The RLS algorithm is summarized as follows [16]:

Parameters:

L_w = filter order

λ = forgetting factor

δ = value to initialize $\mathbf{P}(0) = \begin{cases} \text{small positive constant for high SNR} \\ \text{large positive constant for low SNR} \end{cases}$

Initialization:

$$\begin{aligned}\mathbf{w}(n) &= 0 \\ \mathbf{P}(0) &= \delta^{-1}I,\end{aligned}$$

where I is the $(L_w + 1) \times (L_w + 1)$ identity matrix.

Computation: For $n = 0, 1, 2, \dots$

$$\mathbf{x}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-L_w) \end{bmatrix} \quad (3.47)$$

$$\alpha(n) = d(n) - \mathbf{w}(n-1)^T \mathbf{x}(n) \quad (3.48)$$

$$\mathbf{g}(n) = \mathbf{P}(n-1) \mathbf{x}^*(n) \{ \lambda + \mathbf{x}^T(n) \mathbf{P}(n-1) \mathbf{x}^*(n) \}^{-1} \quad (3.49)$$

$$\mathbf{P}(n) = \lambda^{-1} \mathbf{P}(n-1) - \mathbf{g}(n) \mathbf{x}^T(n) \lambda^{-1} \mathbf{P}(n-1) \quad (3.50)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \alpha(n) \mathbf{g}(n). \quad (3.51)$$

3.1.2.3 Constant Modulus Algorithm.

The Constant Modulus Algorithm (CMA) is a well known and commonly applied blind technique (Bussgang method) developed by Godard [50] and Treichler and Agee [51]. CMA uses the constant modulus criterion in its cost function drives the output signal to a constant amplitude. The cost function is

$$J_{CM}(n) = E \left[(|y(n)|^2 - \gamma)^2 \right]. \quad (3.52)$$

Here γ denotes the dispersion constant and is $\gamma = \frac{E[|x(n)|^4]}{E[|x(n)|^2]^2}$. The error term in (3.30) for the LMS algorithm is replaced with $e(n) = y(n) (\gamma - |y(n)|^2)$. Also, γ is chosen with respect to modulation type. For instance, for the M-ary Phase Shift Keying (M-PSK), $\gamma = 1.0$ and for 16 Quadrature Amplitude Modulation (16-QAM), $\gamma = 1.32$ [40].

Many modulation schemes, such as Gauss Minimum Shift Keying (GMSK), M-PSK and Frequency Modulation (FM), produce waveforms that have a constant envelope. However, the CMA can be used to equalize source constellations that do not have a constant envelope, i.e., 16-QAM [40]. The CMA is robust to carrier-phase offset because the cost function, (3.52), is insensitive to the phase of $y(n)$.

3.1.2.4 Application of the LMS, RLS and CMA Algorithms to the FEQ.

The LMS, RLS and CMA algorithms are expressed in terms of the FEQ. Denote the $NP \times 1$ FEQ vector as

$$\mathbf{D}^T = [\mathbf{D}_1^T, \mathbf{D}_2^T, \dots, \mathbf{D}_P^T]^T \quad (3.53)$$

with p^{th} input vector (FFT output) $\mathbf{u}_p(k) = FFT(\mathbf{y}_p(i))$. Note that for 1 FFT system in Fig. 3.1, $p = 1$ and $\mathbf{u}(k) = FFT(\mathbf{y}(i))$, for 2 FFT system in Fig. 3.2, $p \in 1, 2$ and $\mathbf{u}_1(k) = FFT(\mathbf{y}^1(i))$, $\mathbf{u}_2(k) = FFT(\mathbf{y}^2(i))$. Define an $N \times NP$ FEQ input matrix $\mathbf{U}(k)$:

$$\mathbf{U}(k) = [\text{diag}[\mathbf{u}_1(k)] \mid \text{diag}[\mathbf{u}_2(k)] \mid \dots \mid \text{diag}[\mathbf{u}_P(k)]] \quad (3.54)$$

First, consider the LMS algorithm which is used for initializing the DD-LMS algorithm, where the DD-LMS is used when the training is not available. The $N \times 1$ error vector is $\mathbf{e}(k) = \mathbf{X}(k) - \hat{\mathbf{X}}(k)$, where $\mathbf{X}(k)$ and $\hat{\mathbf{X}}(k) = \mathbf{U}(k)\mathbf{D}$ are the transmitted and estimated signal vectors. The cost function is the norm squared function of this vector:

$$\begin{aligned} J_{mse} &= E[\mathbf{e}^H(k)\mathbf{e}(k)] \\ &= E \left[\left(\mathbf{X}(k) - \hat{\mathbf{X}}(k) \right)^H \left(\mathbf{X}(k) - \hat{\mathbf{X}}(k) \right) \right]. \end{aligned} \quad (3.55)$$

A typical LMS update rule for FEQ coefficients is

$$\mathbf{D}_{lms}(k+1) = \mathbf{D}_{lms}(k) - \mu[\nabla_{\mathbf{D}} J_{mse}], \quad (3.56)$$

where the gradient of the cost function is

$$\nabla_{\mathbf{D}} J_{mse} = -\mathbf{U}^H \left(\mathbf{X}(k) - \hat{\mathbf{X}}(k) \right). \quad (3.57)$$

In defining an RLS algorithm, the cost function is

$$\begin{aligned} J_{mse} = & \mathbf{D}^H E \left[\mathbf{U}^H(k) \mathbf{U}(k) \right] \mathbf{D} \\ & - \mathbf{D}^H E \left[\mathbf{U}^H(k) \mathbf{X}(k) \right] - E \left[\mathbf{X}^H(k) \mathbf{U}(k) \right] \mathbf{D} \\ & + E \left[\mathbf{X}^H(k) \mathbf{X}(k) \right]. \end{aligned} \quad (3.58)$$

With the matrix $\mathbf{R} = E \left[\mathbf{U}^H(k) \mathbf{U}(k) \right]$ and the vector $\mathbf{P} = E \left[\mathbf{U}^H(k) \mathbf{X}(k) \right]$, the optimum FEQ vector is

$$\mathbf{D}_{opt} = \mathbf{R}^{-1} \mathbf{P}. \quad (3.59)$$

Finally, the recursive algorithm is

$$\mathbf{R}(k) = \rho_{rls} \left(\mathbf{R}(k-1) + \mathbf{U}^H(k) \mathbf{U}(k) \right) \quad (3.60)$$

$$\mathbf{P}(k) = \rho_{rls} \left(\mathbf{P}(k-1) + \mathbf{U}^H(k) \mathbf{X}(k) \right) \quad (3.61)$$

$$\mathbf{D}_{rls} = \mathbf{R}^{-1}(k) \mathbf{P}(k), \quad (3.62)$$

where ρ_{rls} is a “forgetting factor” that is slightly less than one (in general). In this recursive algorithm, \mathbf{R} and \mathbf{P} are initialized to all zeros, but the FEQ vector \mathbf{D} is initialized to all ones. Note that the recursion \mathbf{R} is block diagonal and the inverse of \mathbf{R} can be computed easily with Schur complements [52].

These algorithms can be used when training is available, but when it is not decisions can be made on the output data. Then the LMS and the RLS are decision directed (DD):

$$\mathbf{D}_{ddlms}(k+1) = \mathbf{D}_{ddlms}(k) + \mu \mathbf{U}^H \left(Q \left[\hat{\mathbf{X}}(k) \right] - \hat{\mathbf{X}}(k) \right) \quad (3.63)$$

$$\mathbf{P}_{ddrls}(k) = \rho_{rls} \left(\mathbf{P}_{ddrls}(k-1) + \mathbf{U}^H(k) Q \left[\hat{\mathbf{X}}(k) \right] \right), \quad (3.64)$$

where $Q[\cdot]$ quantizes the data to the nearest constellation point.

One of the alternative methods is the constant modulus algorithm. Consider the cost function

$$J_{cma} = E \left[\sum_{i=1}^N \left(\left| \hat{\mathbf{X}}_i(k) \right|^2 - \gamma_i \right)^2 \right] \quad (3.65)$$

where γ_i is a constant representing the modulus of one tone and is chosen as the modulus of the source signal, for example, for QAM constellations, $\gamma = [1, 1, \dots, 1]$. The adaptation rule with respect to CMA cost function is

$$\mathbf{D}_{cma}(k+1) = \mathbf{D}_{cma}(k) - \mu \mathbf{U}^H \left(\left(\hat{\mathbf{X}}^*(k) \odot \hat{\mathbf{X}}(k) - \gamma \right) \odot \hat{\mathbf{X}}(k) \right), \quad (3.66)$$

where \odot stands for element-by-element (Hadamard) multiplication.

3.2 BER Minimizing Channel Shortening with a Genetic Algorithm

The bit error rate minimizing channel shortening equalizer design is not well studied among the CSE design algorithms in the literature. Most CSE designs have attempted to maximize the bit rate. In [12], CSE designs that minimize the BER at the output of the receiver are investigated. Iteratively Reweighted (IR) and Gauss-Newton (GN) Minimum Error Rate (MER) update rules are derived to minimize the multicarrier system, and a greedy search as a heuristic approach for minimizing the SCCP BER model is proposed in [12]. The IR method uses a standard weighted least squares technique to minimize an objective function. However, GN is used to solve non-linear least squares problems. In both algorithms, there are computational complexities, such as taking gradients to minimize the cost function. In this section, we explore the CSE design algorithm in the context of [12] using a Genetic Algorithm, which is simpler than previous methods.

3.2.1 System Model. The MIMO multicarrier system model is shown in Fig. 1.1. In wireless LANs, there are N_a active tones out of N tones, i.e. $N_a = 52$ tones are used to transmit data out of $N = 64$ tones for IEEE 802.11a and HIPERLAN/2 [53, 54]. The set of active tones is \mathcal{S}_a . An N size data symbol is formed by an N_a size finite alphabet data symbols and zero padded to length N . As indicated in the previous chapter, IFFT are taken and a cyclic prefix is inserted in the transmitter. Note that $l \in \{1, \dots, L\}$ is the index of the transmit antenna, $p \in \{1, \dots, P\}$ is the index of the receive antenna, k is the block index, n is the tone index, and i is the sample index.

The received signal is

$$\mathbf{y}(i) = \underbrace{\begin{bmatrix} \mathbf{H}_{1,1}^T & \cdots & \mathbf{H}_{L,1}^T \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1,P}^T & \cdots & \mathbf{H}_{L,P}^T \end{bmatrix}}_{\mathbf{H}^T} \mathbf{x}(i) + \mathbf{n}(i), \quad (3.67)$$

where $\mathbf{x}(i) = [\mathbf{x}_1^T(i), \dots, \mathbf{x}_L^T(i)]^T$ is the transmitted signal vector, $\mathbf{n}(i) = [\mathbf{n}_1^T(i), \dots, \mathbf{n}_P^T(i)]^T$ is the noise vector, and $\mathbf{H}_{l,p}$ is the channel convolution matrix. The output of the CSEs (also the input of the FFT) is

$$\tilde{u}(i) = \mathbf{w}^T \mathbf{y}(i). \quad (3.68)$$

Note that the \mathbf{w} and \mathbf{y} vectors are $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_P^T]^T$ and $\mathbf{y}(i) = [\mathbf{y}_1^T(i), \dots, \mathbf{y}_P^T(i)]^T$. After discarding the CP and taking the FFT, the FEQ input vector is

$$\mathbf{u}(k) = \mathcal{F} \tilde{\mathbf{u}}(k), \quad (3.69)$$

where \mathcal{F} denotes the FFT operation. Assuming that the transmit antenna is $l = 1$, the estimate of the data is the output of the FEQ:

$$\hat{\mathbf{x}}_1(k) = \tilde{\mathbf{d}}_0 \odot \mathbf{u}(k), \quad (3.70)$$

where \odot denotes element-by-element multiplication and $\tilde{\mathbf{d}}_0$ contains the FEQ $\tilde{\mathbf{d}}$, a bank of N_a complex scalars zero-padded to length N .

The SCCP model as in Fig. 1.2 is similar to the multicarrier model. However, since the transmitter operates in the time domain, there are no null tones, $N_a = N$. Keep the same notation as in the multicarrier model, the FEQ output is

$$\hat{\mathbf{u}}(k) = \tilde{\mathbf{d}} \odot \mathbf{u}(k) \quad (3.71)$$

where FEQ vector $\tilde{\mathbf{d}}$ is not zero-padded. Finally, the IFFT output vector is

$$\hat{\mathbf{x}}_1(k) = \mathcal{F}^H \hat{\mathbf{u}}(k). \quad (3.72)$$

3.2.2 BER Models.

3.2.2.1 Multicarrier BER Model. The IEEE 802.11a and HIPER-LAN/2 wireless LAN standards support 4-QAM, 16-QAM and 64-QAM, so the M -level QAM signalling is assumed for the BER computation in this work. Also, it is assumed that the total residual interference and the noise at the output of each tone has a Gaussian PDF. The probability of bit error of the PAM components is [55]

$$P_{\sqrt{M}}(n) = 2 \left(1 - \frac{1}{\sqrt{M}} \right) Q \left(\sqrt{\frac{3}{M-1} \text{SNR}_n} \right), \quad (3.73)$$

and the Symbol Error Rate (SER) on tone n is

$$P_M(n) = 2P_{\sqrt{M}}(n) - (P_{\sqrt{M}}(n))^2, \quad (3.74)$$

where SNR_n is the subchannel SNR (SNR on tone n), and $Q(\cdot)$ is the Q-function. Without loss of generality, the $M = 4$ case is assumed for simplicity of notation, where the BER (3.73) reduces to $Q(\sqrt{\text{SNR}_n})$. The BER and the SER are found by averaging (3.73) and (3.74) over all of the active tones:

$$\text{BER}_{mcm} = \frac{1}{N_a} \sum_{n \in \mathcal{S}_a} Q(\sqrt{\text{SNR}_n}) \quad (3.75)$$

$$\text{SER}_{mcm} = \frac{1}{N_a} \sum_{n \in \mathcal{S}_a} (2Q(\sqrt{\text{SNR}_n}) - Q^2(\sqrt{\text{SNR}_n})) . \quad (3.76)$$

3.2.2.2 Multicarrier Subchannel SNR Model. The subchannel SNR is modeled in different ways in the literature, but most are as in (2.13):

$$\text{SNR}_n = \frac{\mathbf{w}^H \mathbf{B}_n \mathbf{w}}{\mathbf{w}^H \mathbf{A}_n \mathbf{w}}, \quad (3.77)$$

where \mathbf{A}_n and \mathbf{B}_n are Hermitian positive semi-definite matrices.

The FFT output is

$$\underbrace{\mathcal{F}(\mathbf{Y}_k \mathbf{w})}_{\mathbf{u}(k)} = \underbrace{(\mathcal{F} \mathbf{Y}_k)}_{\tilde{\mathbf{Y}}_{k,N}} \mathbf{w}, \quad (3.78)$$

where \mathbf{Y}_k is a $N \times (PL_w)$ block Toeplitz matrix, where each $N \times L_w$ sub-block contains the data that is convolved with the p th CSE \mathbf{w}_p , and successive rows are vectors $\mathbf{y}^T(i)$ for successive values of i . Note that $\tilde{\mathbf{Y}}_{k,N}$ is an $N \times (PL_w)$ matrix with the n th row is $\tilde{\mathbf{y}}_{k,N}$ [12]. The correlation terms are

$$\sigma_n^2 \triangleq \mathcal{E} \{ |\tilde{\mathbf{x}}_1^*(k)[n]|^2 \} \quad (3.79)$$

$$\varphi_n \triangleq \mathcal{E} \{ \tilde{\mathbf{x}}_1^*(k)[n] \tilde{\mathbf{y}}_{k,n} \} \quad (3.80)$$

$$\mathbf{C}_n^2 \triangleq \mathcal{E} \{ \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n} \} \quad (3.81)$$

with dimensions 1×1 , $1 \times PL_w$ and $PL_w \times PL_w$, [12].

The subchannel SNR is measured at the output of the FEQ, where the desired signal is an undistorted copy of the transmitted signal. Also, the error is total interference and noise for block k and tone n , i.e. $\tilde{e}_{k,n} = \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{\mathbf{x}}_1(k)[n]$. Thus, the subchannel SNR on tone n is the ratio of the power of the desired signal to the power of the error [12]:

$$\text{SNR}_n = \frac{\sigma_n^2}{\mathcal{E}\{|\tilde{e}_{k,n}|^2\}} = \frac{\sigma_n^2}{\mathcal{E}\{|\tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{\mathbf{x}}_1(k)[n]|^2\}} \quad (3.82)$$

$$= \frac{\sigma_n^2}{|\tilde{d}_n|^2 \mathbf{w}^H \mathbf{C}_n^2 \mathbf{w} - \tilde{d}_n \varphi_n \mathbf{w} - \tilde{d}_n^* \varphi_n^* \mathbf{w}^* + \sigma_n^2}. \quad (3.83)$$

If the correlation of the input and output on tone n are set equal to the transmitted power on tone n , the unbiased MMSE FEQ is [12]

$$\tilde{d}_n^{u\text{MMSE}} = \frac{\sigma_n^2}{\varphi_n \mathbf{w}}. \quad (3.84)$$

Substituting the (3.84) into (3.83) yields

$$\text{SNR}_n = \frac{1}{\sigma_n^2 \left(\frac{\mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}}{\mathbf{w}^H \varphi_n^H \varphi_n \mathbf{w}} \right) - 1} \quad (3.85)$$

$$= \frac{\mathbf{w}^H \mathbf{B}_n \mathbf{w}}{\mathbf{w}^H \mathbf{A}_n \mathbf{w}} \quad (3.86)$$

with

$$\mathbf{A}_n = \sigma_n^2 \mathbf{C}_n^2 - \varphi_n^H \varphi_n \quad (3.87)$$

$$\mathbf{B}_n = \varphi_n^H \varphi_n. \quad (3.88)$$

3.2.2.3 SCCP BER Model. Due to the lack of null tones in SCCP, i.e., $N_a = N$, the SCCP BER is averaged over all N samples of the IFFT output. Under the same assumptions as in the multicarrier model, the probability of error on

the PAM component of sample m is [55]

$$P_{\sqrt{M}}(m) = 2 \left(1 - \frac{1}{\sqrt{M}} \right) Q \left(\sqrt{\frac{3}{M-1} \text{SNR}_m} \right), \quad (3.89)$$

and the SER of sample m is

$$P_M(m) = 2P_{\sqrt{M}}(m) - (P_{\sqrt{M}}(m))^2, \quad (3.90)$$

where SNR_m is the effective signal-to-interference and noise ratio on sample m . The BER and SER for $M = 4$ with averaging (3.89) and (3.90) over the N samples are

$$\text{BER}_{sccp} = \frac{1}{N} \sum_{m=1}^N Q(\sqrt{\text{SNR}_m}) \quad (3.91)$$

$$\text{SER}_{sccp} = \frac{1}{N} \sum_{m=1}^N (2Q(\sqrt{\text{SNR}_m}) - Q^2(\sqrt{\text{SNR}_m})) . \quad (3.92)$$

3.2.2.4 SCCP Output SNR Model. The FFT output is as in (3.78) (as in multicarrier model). But the FEQ $\tilde{\mathbf{d}}$ output is different:

$$\hat{\mathbf{x}}_1(k)[m] = \sum_{n=1}^N \mathbf{Q}_{m,n} \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w}, \quad (3.93)$$

where $\mathbf{Q}_{m,n}$ is the $m \times n$ (unitary) IFFT matrix $\mathbf{Q} = \mathcal{F}^H$. The correlation terms are

$$\sigma_m^2 \triangleq \mathcal{E} \{ |\tilde{\mathbf{x}}_1^*(k)[m]|^2 \} \quad (3.94)$$

$$\varphi_{m,n} \triangleq \mathcal{E} \{ \tilde{\mathbf{x}}_1^*(k)[m] \tilde{\mathbf{y}}_{k,n} \} \quad (3.95)$$

$$\mathbf{C}_{m,n}^2 \triangleq \mathcal{E} \{ \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n} \} \quad (3.96)$$

with dimensions the same as before. Assume the N transmit samples have identical power, so σ^2 is independent of m . The output SNR of sample m is the ratio of the desired signal power to the error power, i.e., $\tilde{e}_{k,m} = \hat{\mathbf{x}}_1(k)[m] - \tilde{\mathbf{x}}_1(k)[m]$. Hence, the

output SNR is

$$\text{SNR}_m = \frac{\sigma^2}{\mathcal{E}\{|\tilde{e}_{k,m}|^2\}} = \frac{\sigma^2}{\mathcal{E}\left\{\left|\sum_{n=1}^N \mathbf{Q}_{m,n} \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{\mathbf{x}}_1(k)[m]\right|^2\right\}}. \quad (3.97)$$

Expanding the denominator yields

$$\begin{aligned} \mathcal{E}\{|\tilde{e}_{k,m}|^2\} &= \sum_{n_1, n_2} \mathbf{Q}_{m,n_1}^* \mathbf{Q}_{m,n_2} \tilde{d}_{n_1}^* \tilde{d}_{n_2} \mathbf{w}^H \mathbf{C}_{n_1, n_2}^2 \mathbf{w} - \sum_n \mathbf{Q}_{m,n} \tilde{d}_n \varphi_{m,n} \mathbf{w} \\ &\quad + \sum_n \mathbf{Q}_{m,n}^* \tilde{d}_n^* \varphi_{m,n}^* \mathbf{w}^* + \sigma^2, \quad m \in \{1, \dots, N\}. \end{aligned} \quad (3.98)$$

As in the multicarrier model, the unbiased MMSE FEQ for sample m in SCCP is

$$\mathcal{E}\left\{\tilde{\mathbf{x}}_1^*(k)[m] \sum_{n=1}^N \mathbf{Q}_{m,n} \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w}\right\} = \mathcal{E}\{\tilde{\mathbf{x}}_1^*(k)[m] \tilde{\mathbf{x}}_1(k)[m]\} \quad (3.99)$$

$$\sum_{n=1}^N \mathbf{Q}_{m,n} \tilde{d}_n \varphi_{m,n} \mathbf{w} = \sigma^2, \quad (3.100)$$

Finally, the output SNR is [12]

$$\text{SNR}_m = \frac{\sigma^2}{\sum_{n_1, n_2} \mathbf{Q}_{m,n_1}^* \mathbf{Q}_{m,n_2} \tilde{d}_{n_1}^* \tilde{d}_{n_2} \mathbf{w}^H \mathbf{C}_{n_1, n_2}^2 \mathbf{w} - \sigma^2}. \quad (3.101)$$

Also, (3.100) in matrix form is

$$\tilde{\mathbf{d}}^T \mathbf{A}_m \mathbf{w} = \sigma^2, \quad m \in \{1, \dots, N\}, \quad (3.102)$$

where row $\mathbf{A}_m[n, :] = \mathbf{Q}_{m,n} \varphi_{m,n}$. These N equations are gathered into a vector and solved for $\tilde{\mathbf{d}}$ [12]:

$$\tilde{\mathbf{d}} = \sigma^2 \begin{bmatrix} \mathbf{w}^T \mathbf{A}_1^T \\ \vdots \\ \mathbf{w}^T \mathbf{A}_N^T \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \quad (3.103)$$

3.2.3 Applying Genetic Algorithm. The IR and GN-MER algorithms proposed in [12] have some complexities. One is taking derivatives with respect to some vectors, i.e., \mathbf{w} and \mathbf{d} . Another is Lagrangian analysis of the cost function. Also, gradient analysis of the SCCP BER model is mathematically intractable.

For these reasons, the Genetic Algorithm (GA) is proposed, which is simpler than the discussed algorithms, to minimize the BER model of the multicarrier and SCCP systems. As indicated in Chapter 2, the GA does not require derivative information, and it optimizes variables with extremely complex cost surfaces and searches from a wide sampling of the cost surface.

In Chapter 2, the GA is considered in general. In this section, the application of the GA to the proposed CSE technique is considered. The GA used here is a continuous GA [48].

First, a chromosome of the GA must be determined. Due to searching for the CSE coefficients which give minimum BER, the CSE coefficients are selected as a chromosome. A chromosome is formed by the CSE coefficients of size L_w :

$$chromosome = \underbrace{\begin{bmatrix} c_{w(0)} & c_{w(1)} & \cdots & c_{w(L_w-2)} & c_{w(L_w-1)} \end{bmatrix}}_{L_w}.$$

After determining the chromosome, an initial population is gathered with many different chromosomes. In order to avoid searching the whole space, the population is not initialized with random numbers but is formed in the neighborhoods of the Maximum Shortening Signal-to-Noise Ratio (MSSNR) [26]/minimum Interblock Interference (Min-IBI) [32] designs. A step size $([-0.1 \ 0.1])$ is selected to explore this neighborhood. The population size is decided by the user (in Chapter 4, population size is selected as 10).

The costs of each chromosome are calculated using equation (3.75) for a multicarrier system and (3.91) for an SCCP system. According to the principle of the survival of the fittest, the five fittest (this number considers half of the population

Table 3.1: Rank weighting in a GA.

n	Chromosome	P_n	$\sum_{i=1}^N P_i$
1	<i>chromosome₁</i>	0.3333	0.3333
2	<i>chromosome₂</i>	0.2667	0.6
3	<i>chromosome₃</i>	0.2	0.8
4	<i>chromosome₄</i>	0.1333	0.9333
5	<i>chromosome₅</i>	0.0667	1

and can be changed by the user) which give the minimum cost are kept and the others are discarded from the population.

These (better) chromosomes are called parents and are mated to generate the new generation. However, the selection method for pairing the chromosomes can be chosen at this point. *Rank weighting* is used within the *weighted random pairing* methods. In this method, the chromosomes are sorted from lowest to the highest cost and the selection probability of a rank n chromosome is found as

$$P_n = \frac{N_{keep} - n + 1}{\sum_{n=1}^{N_{keep}} n}, \quad (3.104)$$

where N_{keep} is the number of chromosomes that are kept inside the population. As shown in Table 3.1, the probability of selecting the lowest cost chromosome is higher.

Two children are generated from every parent. In doing this, a crossover point is randomly selected and, the variables to/from this point are exchanged. A next generation is formed as follows:

1. Two parents that mate are selected as described above from the population.
2. A crossover point between the first and last digit of a chromosome is randomly selected.
3. The first part of the *parent₁* and the second part of the *parent₂* are combined to generate the first child.

Table 3.2: An example of mutation and new population.

Population after mating	Cost	Population after mutations	New cost
[0.21 -0.121 0.002 ... -0.88]	0.1245	[0.21 -0.121 0.002 ... -0.88]	0.1245
[0.45 -0.09 -0.18 ... -0.03]	0.1341	[0.45 -0.09 -0.18 ... -0.03]	0.1341
[-0.62 0.19 0.11 ... 0.22]	0.2145	[-0.62 <u>0.23</u> 0.11 ... 0.22]	0.2167
[-0.34 0.74 0.32 ... -0.11]	0.3267	[-0.34 <u>0.14</u> 0.32 ... <u>-0.67</u>]	0.3741
[0.99 0.1 0.01 ... 0.52]	0.3566	[<u>0.09</u> <u>0.21</u> 0.01 ... <u>0.72</u>]	0.1127
[0.21 0.801 -0.118 ... -0.103]	0.4132	[0.21 0.801 -0.118 ... -0.103]	0.4132
[0.43 0.88 -0.77 ... -0.221]	0.4629	[0.43 <u>-0.288</u> -0.77 ... -0.221]	0.5288

4. The second part of the $parent_1$ and the first part of the $parent_2$ are combined to generate the second child.

After generating the new chromosomes, population size attains the initial size. At this point, mutation is applied to make a detailed search of the cost function and to avoid converging to local minima. In order to implement mutation, a randomly selected value in a randomly selected chromosome is changed to a new value within the step size:

$$[0.23 \quad \underline{0.012} \quad -0.56 \quad \dots \quad -0.023] \implies [0.23 \quad \underline{\mathbf{0.107}} \quad -0.56 \quad \dots \quad -0.023]$$

The final population is formed by the new chromosomes, and the cost of each chromosome is calculated after mutation (see Table 3.2). The algorithm then returns to the beginning to implement the next generation. The number of generations depends on whether an acceptable minimum is reached or a set number of iterations is exceeded. Here, the iteration criteria to stop the algorithm is selected. (In Chapter 4, the algorithm is stopped after 40 iterations.)

IV. Simulations and Results

4.1 Results for Post-FEQ Diversity Combining

In this section, a multicarrier system with FFT size $N = 64$ and a CP length of $v = 16$ is considered. This work can be compared to the IEEE 802.11a and HIPERLAN standards, but no particular standard is implemented. The SNR in the simulations is 20 dB, and there are $P = 4$ receive antennas. Each channel is modeled as a 32-tap Rayleigh-fading channel with an exponential delay profile and with approximately 32 nonzero taps (about double the CP length). The length of each CSE is $L_w = 40$ taps and the CSE/CSEs that has/have a unit tap is $p = 1$ for 1 FFT system, $p = 1$ and $p = 3$ for 2 FFT system and each CSE has a unit tap for 4 FFT system. In every CSE, the $tap = \lfloor (L_w/2) - 1 \rfloor = 19$ th tap (selected arbitrarily) is set to one. The forgetting factor for RLS-MERRY is $\rho = 1$, and the step size for LMS and CMA algorithms and for adaptive MMSE is $\mu = 0.05$ and $\mu = 10^{-4}$.

There are bounds on μ , step size parameter of the LMS algorithm, in the literature. Considering the correlation matrix of input signals, Eqn. (3.27), the step size is chosen as

$$0 < \mu < \frac{2}{\lambda_{max}}, \quad (4.1)$$

where λ_{max} is the largest eigenvalue of \mathbf{R} [16].

The simulations have two parts. The first part has two CSE adaptation algorithms as RLS-MERRY and Adaptive MMSE (A-MMSE). The second part has the FEQ adapt with CMA, DD-LMS and DD-RLS algorithms for RLS-MERRY and RLS for A-MMSE. Fig. 4.1 shows MERRY cost with RLS-like CSE adaptation algorithm for the old architecture and the two new architectures. The cost is measured as in (3.13), where outputs can be computed in (3.3), (3.4) and (3.5). These plots show that MERRY is working if the cost functions decrease, but the relative values of the cost functions are not significant.

Fig. 4.2 shows BER performance of the proposed post combining techniques as the FEQ adapts. To obtain these figures, the FEQ is adapted after the CSE has

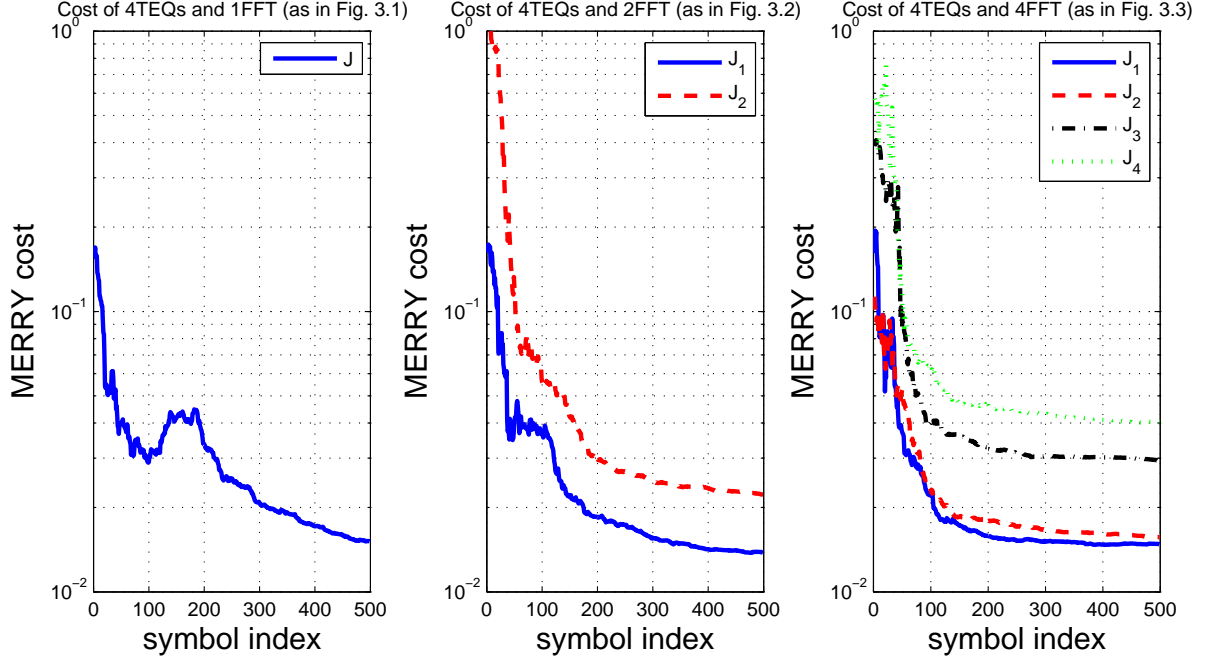


Figure 4.1: MERRY Cost vs. CSE iteration. The first sub-figure (left) shows the system which has 1 *FFT*, second shows 2 *FFT*, and third shows P *FFT* (in this simulation $P = 4$).

converged. These algorithms are compared with adaptive MMSE in [17] for reference. In the figure, for instance, RLS-MERRY/CMA indicates that the CSE adaptation algorithm is RLS-MERRY and FEQ adaptation algorithm is CMA, etc.

As stated in [49] the blind implementations of these algorithms usually do not converge to adequate FEQ taps. Thus, the semi-blind technique is considered in simulations which are initializations of DD-LMS and CMA algorithms with a small number of iterations of LMS, and DD-RLS with a small number of iterations of RLS. The DD-LMS and CMA algorithms are initialized by 25 iterations of LMS, and the DD-RLS are initialized by 5 iterations of RLS.

Examining the Fig. 4.2 indicates that the RLS algorithm converges faster and has $\text{BER} \approx 10^{-2}$. However, the 2 *FFT* system in Fig. 3.2 shows better performance, i.e., $\approx 2.5 \times 10^{-3}$ BER, but the P *FFT* system in Fig. 3.3 shows the best performance, i.e., a lower BER $\approx 1.5 \times 10^{-4}$. Hence, there is a two order of magnitude improvement,

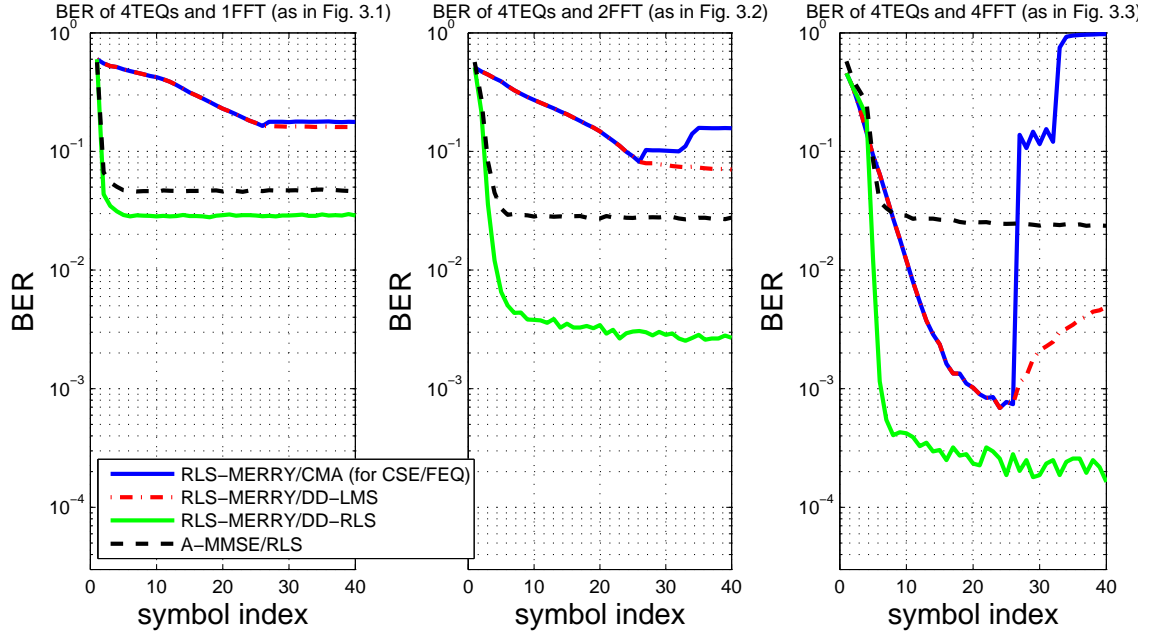


Figure 4.2: BER vs. FEQ iteration after the CSE has converged. The left subfigure shows the system which has 1 *FFT*, middle shows 2 *FFT* and right shows P *FFT* (in this simulation $P = 4$).

i.e., 20 dB, in BER. Combining the diversity after the FEQ shows better performance than combining after the CSE.

4.2 Results for Channel Shortening with GA

In this section, the proposed CSE design is simulated with GA and compared to several designs in the literature. The compared algorithms are the MSSNR method, the MMSE method, the MDS method, the min-IBI method, the Gauss-Newton implementation of bit rate maximizing design (GN-BM), the IR-MER method, the GN-MER method and the greedy minimum error rate (G-MER) method. The simulations are also compared to the matched filter bound (MFB) and the BER when no CSE is used.

4.2.1 Results of GA for Multicarrier Systems. Again a wireless system with $N = 64$ FFT size, $v = 16$ CP length, and $N_a = 52$ active tones (12 null tones) as in

wireless LAN standards, i.e. IEEE 802.11a, HIPERLAN/2 is considered. The active tones use 4-QAM constellations and the channels are 32 tap Rayleigh fading channels with exponential delay profiles. There are $L = 1$ transmit antenna and $P = 2$ receive antennas. The CSEs have $L_w = 16$ taps. The correlation parameters σ^2, φ_n and \mathbf{C}_n^2 are estimated using 2000 symbols of training, and the iterative algorithms (IR-MER, GN-MER, GN-BM) run for 40 iterations.

The GA has 10 population size, 0.15 mutation rate and 50% of a population is kept for the next generation. A chromosome is formed by CSE taps, so it has length 32.

The computational complexity is explained in terms of the number of complex multiply-and-accumulate (MAC) operations. The IR-MER, GN-MER and GA have $(1/2)(PL_w)^2 N_a N_{corr}$ complex MACs to compute correlations in which N_{corr} symbols are used to compute the correlation terms. The IR-MER and GN-MER algorithms require further $(3/2)(PL_w)^2 N_a + (1/3)(PL_w)^3$ and $(PL_w)^2 N_a + 4(PL_w) N_a + (1/3)(PL_w)^3$ complex MACs per iteration respectively [12]. However, GA requires further $(PL_w) N_a N_{pop}$ complex MACs per generation. Because the IR-MER and GN-MER have 40 iterations in [12], the population and generation size are chosen as $N_{pop} = 10$ and $N_{gen} = 7$, which gives the GA the same complexity as IR-MER and GN-MER.

The results are averaged over 500 independent channels, and input and noise sequences. The BER is measured over 1000 symbols for 0-20 dB SNR, 2000 symbols for 25 dB SNR and 4000 for 30 dB SNR for each channel realization. The MSSNR design is used as the initialization for the iterative algorithms and the GA.

Figs. 4.3 and 4.4 show the BER vs. SNR and SER vs. SNR in dB. Below 15 dB there is no need to use MSSNR design compared to no CSE. In order to provide same complexity with GN-MER and IR-MER, GA uses 10 population size and 7 generation. This is far away from converging to a global minimum. Although the GA does not have enough generations, it has still better performance than MSSNR

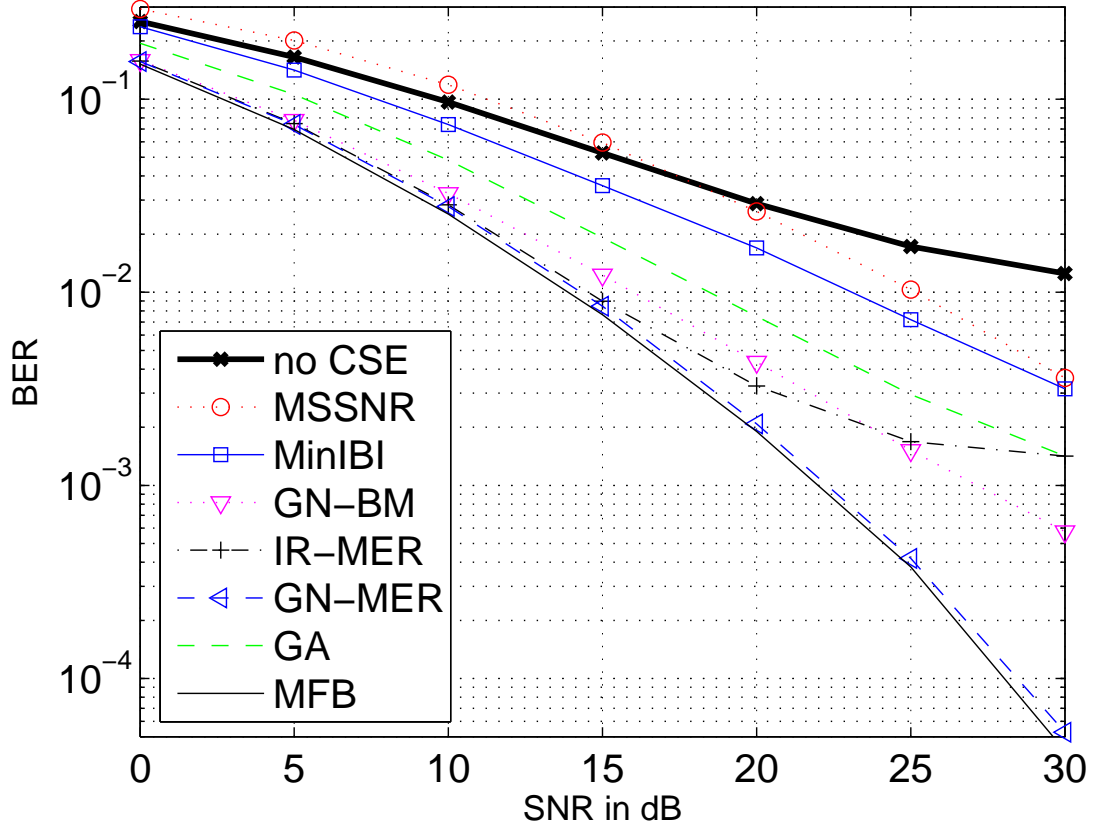


Figure 4.3: BER vs. SNR for a multicarrier system when no CSE, MSSNR, min-IBI, GN-BM, IR-MER, GN-MER and the proposed GA is used. There are two receive antennas, the channels are Rayleigh fading with 32 taps each, CSE has 16 taps, and there are 52 active tones and 12 null tones. The CP length is 16 and all of the iterative algorithms use 40 iterations. The GA uses 10 population and 7 generation to provide the same complexity.

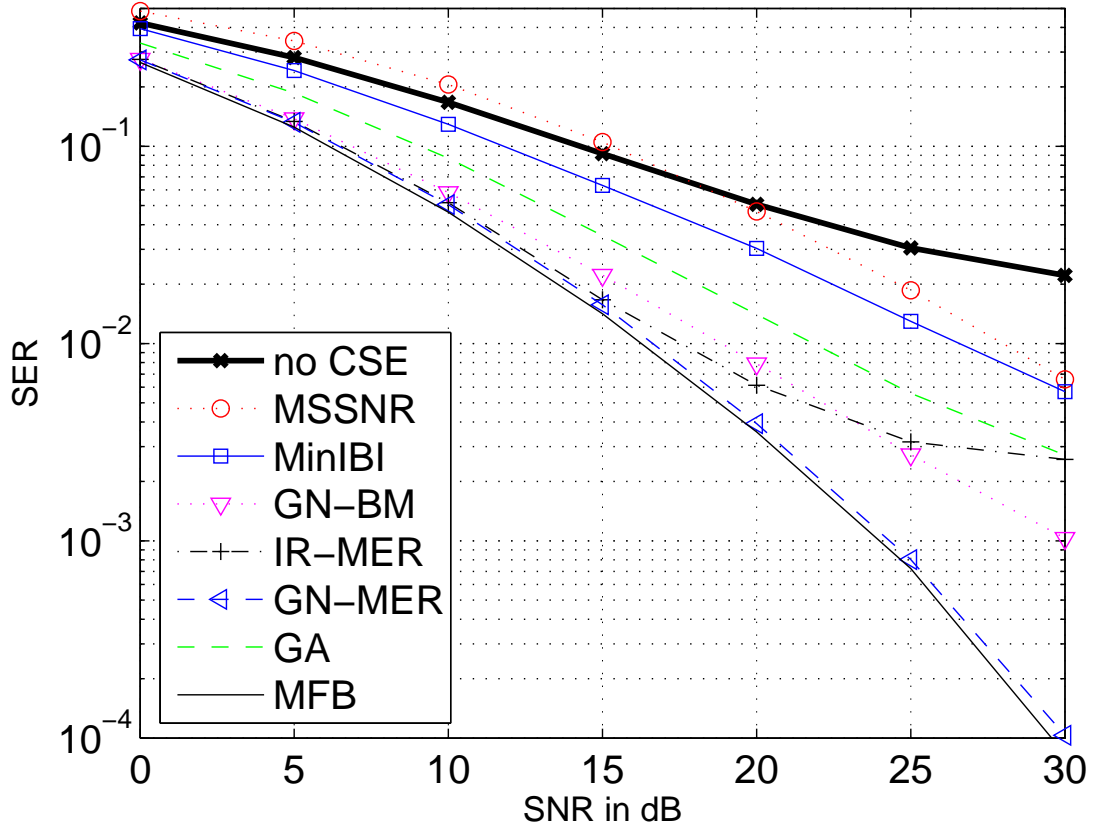


Figure 4.4: SER vs. SNR for a multicarrier system. All of the conditions are same as in Fig. 4.3

and Min-IBI. So, in general there is no need to use GA, which has same complexity as IR-MER and GN-MER. However, in the next section, it is shown that GA does provide an improvement in SCCP systems.

4.2.2 Results of GA for SCCP Systems. The SCCP system with $N = 16$ FFT size, $v = 4$ CP length is considered. These parameters are chosen to be consistent with [12]. The Rayleigh fading channels have 10 taps with exponential delay profiles. Again, there are $L = 1$ transmit antennas and $P = 2$ receive antennas. Each CSE has $L_w = 16$ taps. The GA parameters are same as for the multicarrier system.

The correlation terms, σ^2 , φ_n and \mathbf{C}_n^2 , are estimated using 100 blocks of training, and the min-IBI design is selected for initialization of the greedy search and the GA.

The BER is measured over 250 independent channels, with input data and noise sequences using 2000 blocks of data each.

Complexity is computed to determine the GA population and generation size. The greedy search and GA require $(1/2)(PL_w)^2 N^2 N_{corr}$ complex MACs to compute correlation terms. Additionally, the greedy search requires a further $(PL_w)^2 N^3$ complex MACs per iteration [12]. The GA, on the other hand, requires $(PL_w)^2 N^3 N_{pop}$ per generation. Because the greedy search has 400 iterations in [12], the population and generation size are chosen as $N_{pop} \times N_{gen} = 400$, for example, $N_{pop} = 10$ and $N_{gen} = 40$.

Fig. 4.5 shows the BER vs. SNR for SCCP. Below 15 dB there is no need to use MDS and MSSNR designs compared to no CSE. In order to provide the same complexity with G-MER, the GA uses 10 population size and 40 generations which is not enough to converge a global minimum. However, below 10 dB GA provides relatively the same performance as G-MER.

Since some error correction codes operate on symbols rather than bits, the SER is a better metric than the BER. Then the results for SER show more accurate information about the performance of GA. Fig. 4.6 shows SER vs. SNR. Although it does not have enough population and generation sizes, GA provides better performance below 17 dB.

When the complexity is not a consideration, the GA can find a better solution. Figs. 4.7 & 4.8 show results for GA for more populations and generations. The population and generation sizes are selected as $N_{pop} = 20$ and $N_{gen} = 200$ for these results. Both the BER and the SER performances are much better than G-MER. Approximately a factor of 3 magnitude improvement in BER and a factor of 4 magnitude improvement in SER are obtained with GA. Also, note that the G-MER algorithm can not find a better solution than the results in figures even if the iteration number is increased.

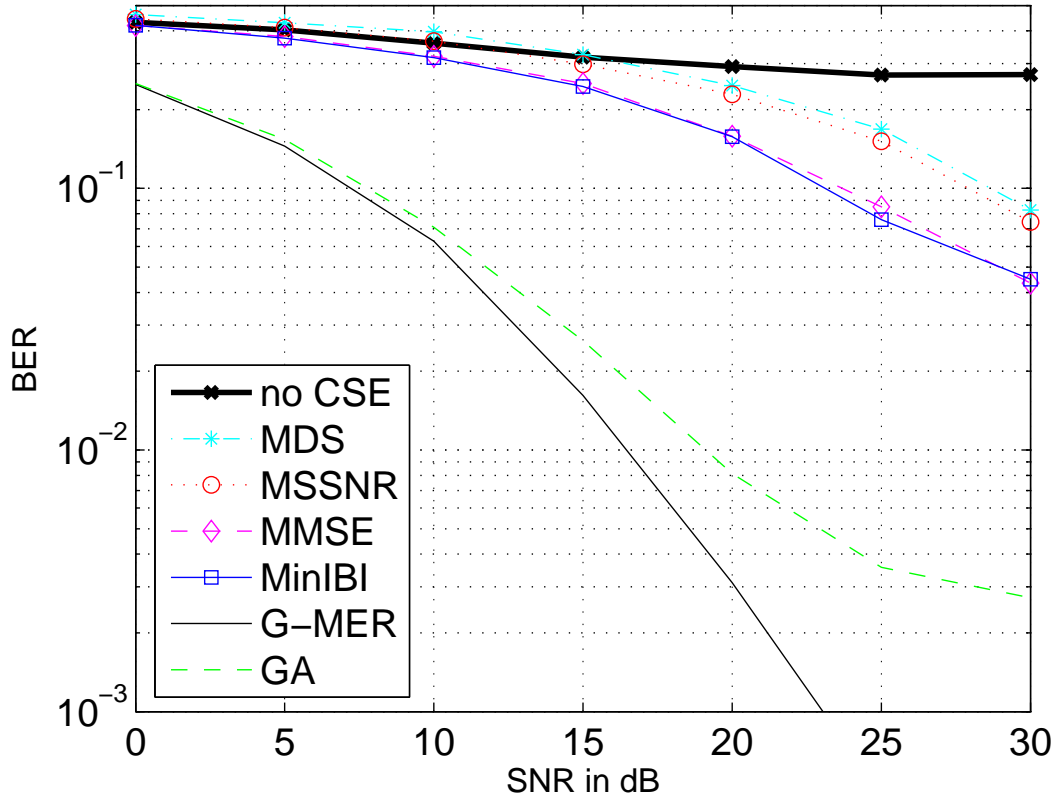


Figure 4.5: BER vs. SNR for SCCP system when no CSE, the MDS, the MSSNR, the min-IBI, the MMSE, the G-MER and the proposed GA are used. There are two receive antennas, the channels are Rayleigh fading with 10 taps each, CSE has 16 taps and the FFT size is 16. CP length is 4 and all of the iterative algorithms use 400 iterations, and the GA uses 10 populations and 40 generations to provide the same complexity.

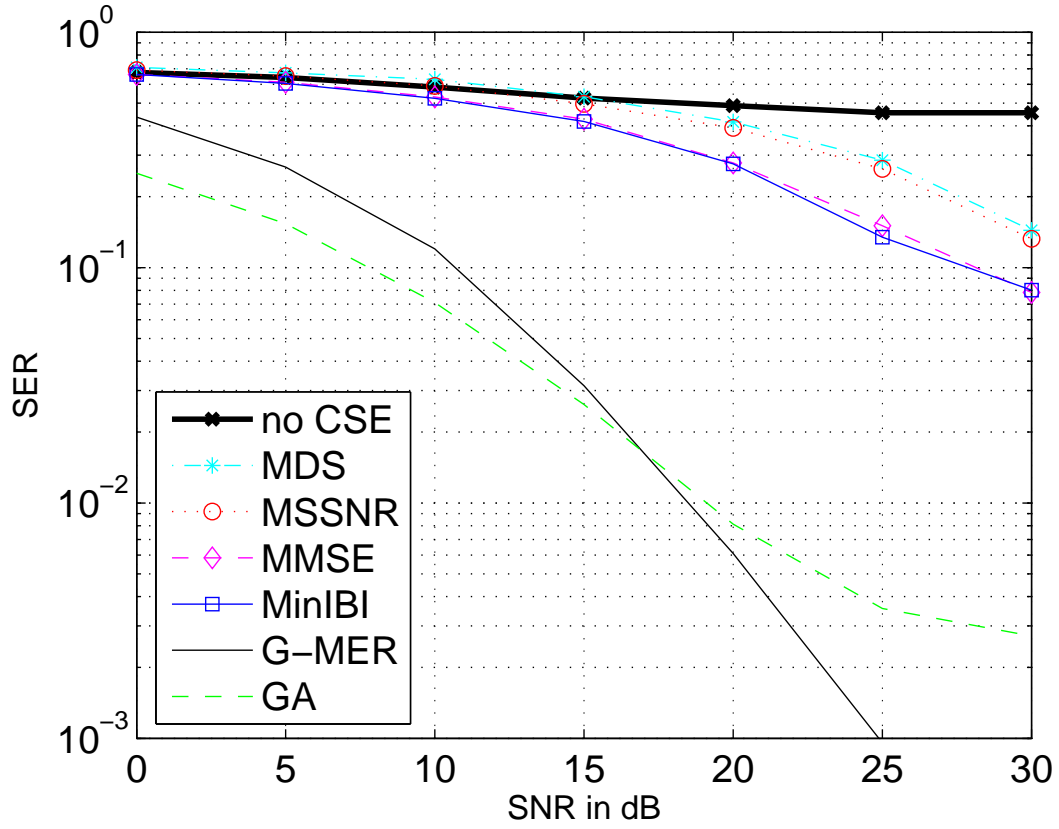


Figure 4.6: SER vs. SNR for SCCP system. All of the conditions are same as in Fig. 4.5

Finally, note that GA is a very efficient search method for multimodal cost functions.

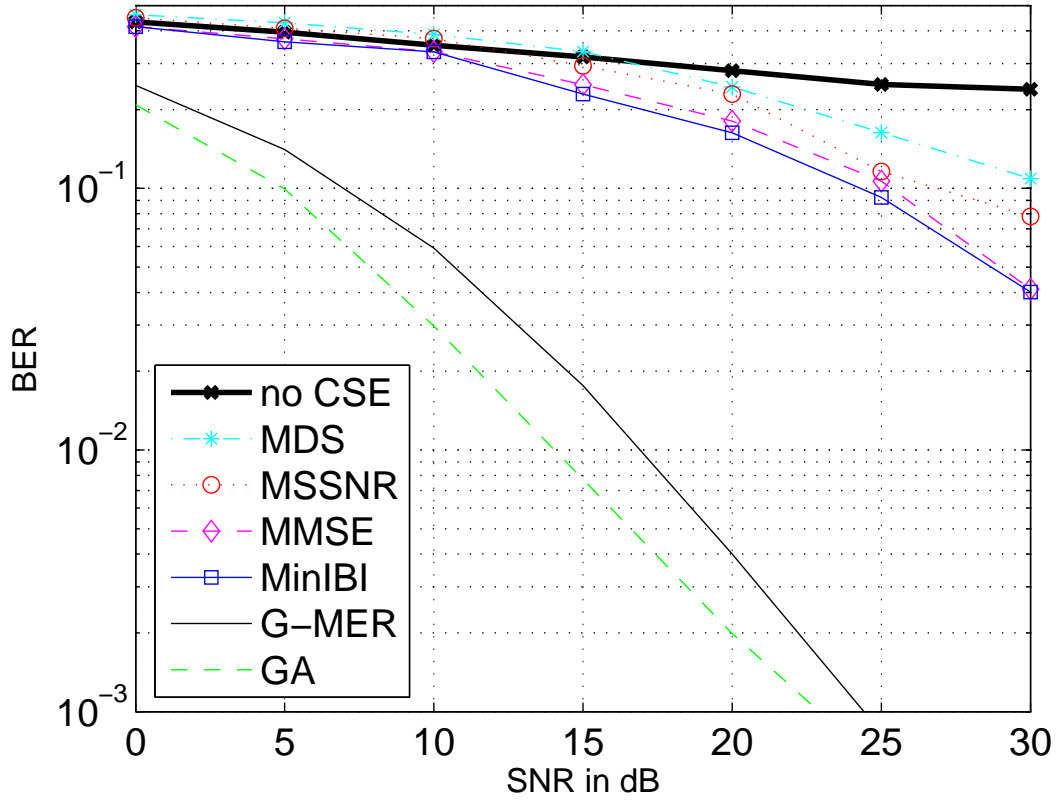


Figure 4.7: BER vs. SNR for SCCP system when no CSE, the MDS, the MSSNR, the min-IBI, the MMSE, the G-MER and proposed GA are used. There are two receive antennas, the channels are Rayleigh fading with 10 taps each, CSE has 16 taps and the FFT size is 16. CP length is 4 and all of the iterative algorithms use 400 iterations, and the GA uses 20 populations and 200 generations.

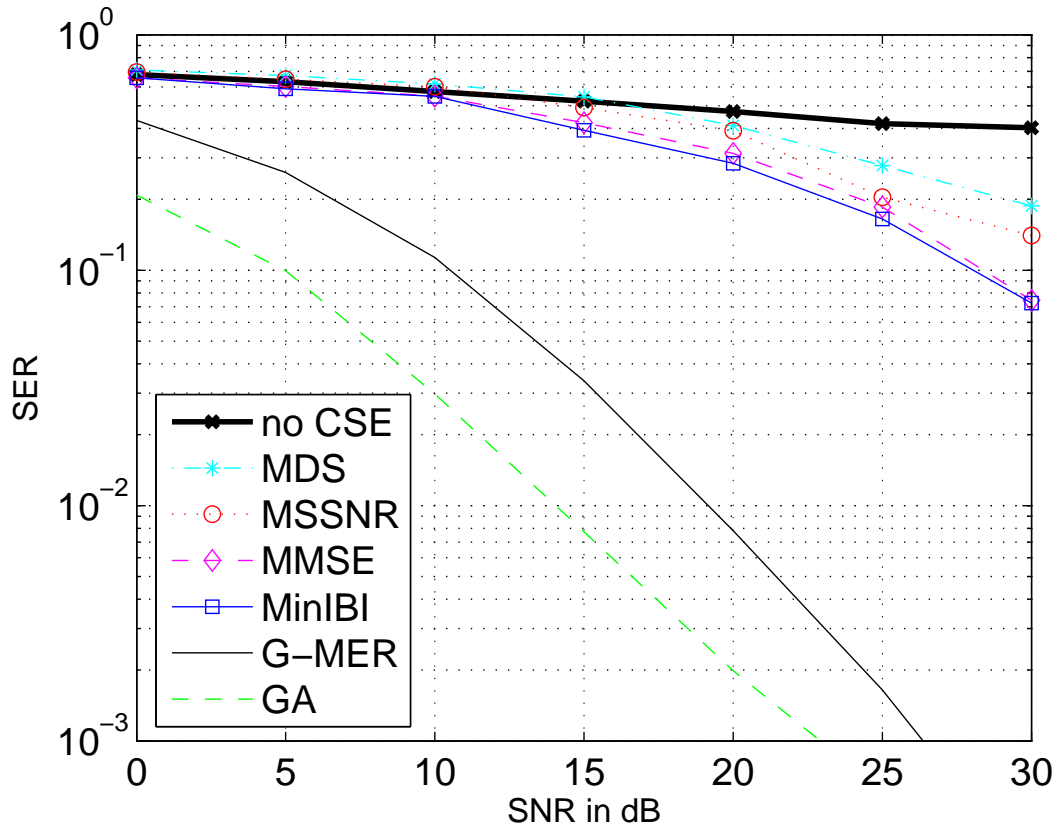


Figure 4.8: SER vs. SNR for SCCP system. All of the conditions are same as in Fig. 4.7

V. Conclusions

MULTIMEDIA applications and the development of the Internet require high speed and wide band digital communications. Inter-Symbol Interference (ISI) is one of the main problems in these broadband channels. Multicarrier modulation (MCM) is used to prevent ISI. Examples of wireline MCM systems include power line communications (HomePlug) and digital subscriber lines (DSLs) [56], and examples of wireless systems include wireless local area network (IEEE 802.11a/g, HIPERLAN/2, MMAC) [7], wireless metropolitan area networks (IEEE 802.16) [8], digital video and audio broadcasting in Europe [9, 10], satellite radio (Sirius and XM Radio) [11], and the proposed standard for multiband ultra wideband (IEEE 802.15.3a).

In MCM, a guard period which is the last v samples of the MCM frame is added in front of the MCM frame. This type of guard period is called Cyclic Prefix (CP). If the channel is short, equalization can be done by a Frequency Domain Equalizer (FEQ), which is a bank of scalars. However, if the channel is longer than the CP, an equalizer called a Channel Shortening Equalizer (CSE) which shortens the channel impulse response to the length v or less, is added to the receiver front end.

Most channel shortener design methods in the literature assume knowledge of channel state information (CSI). This information can be obtained by estimating the channel. Therefore, a training sequence is sent with the transmitted signal, but this training reduces the channel capacity. Also, in wireless systems the channel changes rapidly, and the training must be sent more frequently. Hence, there are blind and semi-blind methods in channel equalization techniques.

Early channel shortening techniques were based on heuristic objective functions. In recent algorithms, the goal is maximizing the bit rate for a given Bit Error Rate (BER), which is a proper performance measurement for wireline multicarrier systems, e.g., DSL. In contrast, wireless multicarrier, e.g., OFDM, and SCCP systems have a fixed bit loading algorithm, and the performance metric is the BER for a fixed bit rate. Additionally, the CSE is updated to minimize the BER for the initial bit loading even in DSL.

The main objective of this thesis is to design a CSE and to modify the system architecture to obtain a lower BER. This thesis has two parts: one explores modifying an OFDM system architecture and adapting the FEQ to obtain a lower BER; another proposes to design a CSE adapted with a Genetic Algorithm (GA) to minimize BER.

In the first part, the system architecture is modified to obtain a lower BER in the context of [49]. The recursive least squares (RLS)-like implementation of the Multicarrier Equalization by Restoration of RedundancY (MERRY) channel shortener is used to shorten the channel. Three types of adaptation rules, i.e., DD-Least Mean Squared (LMS), DD-Recursive Least Squares (RLS), and Constant Modulus Algorithm (CMA), are derived for the new architectures and are used to adapt the FEQs. To minimize BER, different types of diversity combining are considered.

The proposed architectures can be classified as 1 FFT, 2 FFT and P FFT systems. The 1 FFT system is the traditional OFDM system architecture. The 2 FFT system selects the diversity point after the FFT/FEQ instead of before the FFT. The P FFT system generalizes the proposed technique. The idea of the architecture design is to restore the values of the tones affected by deep fades.

Simulation results show that the second proposed technique (P FFT system) has best performance, i.e., a lower BER. There is a two order of magnitude improvement, i.e., 20 dB, in BER with the modified system architecture. Thus, combining the diversity after the FEQ yields better performance than combining after the CSE.

In the second part, a CSE for multicarrier and Single Carrier Cyclic Prefixed (SCCP) systems is designed by adapting with GA, which is a global search method based on the principles of natural selection and genetics. The GA fitness function is the BER. An initial population is formed by a maximum shortening signal-to-noise ratio (MSSNR) design for the multicarrier system and by a minimum inter block interference (min-IBI) design for the SCCP system. Mutation is used, which makes a detailed search of the cost function and avoids convergence to local minima.

The complexities of the algorithms proposed in [12] and GA are compared in order to design CSE with comparable complexity but better performance. The simulation results show that the required small N_{pop} and N_{gen} sizes do not indicate that the adaptation of the CSE for multicarrier systems yield better performance than the algorithms in [12]. However, GA shows better performance for symbol error rate (SER) in SCCP system between 0 and 17 dB, even though it does not have enough population and generation size to fully converge.

However, the GA can find an even better solution, i.e., a factor of 3 magnitude improvement in BER and a factor of 4 magnitude improvement in SER, in a SCCP system, when increased complexity is not avoided.

Future work can focus on designing an algorithm which has lower complexity or shows better performance than GA.

Bibliography

1. G. Arslan, "Equalization of discrete multitone transceivers," Ph.D. dissertation, University of Texas at Austin, December 2000.
2. R. K. Martin *et al.*, "Unification and evaluation of equalization structures and design algorithms for discrete multitone modulation systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3880–3894, October 2005.
3. A. Duel-Hallen and C. Heegard, "Delayed decision feedback sequence estimation," *IEEE Trans. Comm.*, vol. 37, pp. 428–436, May 1989.
4. M. V. Eyuboglu and S. U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 16–20, January 1988.
5. M. Milošević, "Maximizing data rate of discrete multitone systems using time domain equalization design," Ph.D. dissertation, University of Texas at Austin, May 2003.
6. M. Ding, "Channel equalization to achieve high bit rates in discrete multitone systems," Ph.D. dissertation, University of Texas at Austin, August 2004.
7. R. D. J Van Nee *et al.*, "New high rate wireless LAN standards," *IEEE Commun. Mag.*, vol. 37, no. 12, pp. 82–88, December 1999.
8. IEEE Std. 802.16a, "Air interface for fixed broadband wireless access systems, MAC and additional PHY specifications for 2–11 GHz," 2003.
9. ETSI EN 300 744 V1.4.1, The European Telecomm. Standards Inst., "Digital Video Broadcasting (DVB): framing structure, channel coding and modulation for digital terrestrial telev." 2003.
10. ETSI EN 300 401, The European Telecomm. Standards Inst., "Radio broadcasting system, Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers," 2001.
11. D. H. Layer, "Digital radio takes to the road," *IEEE Spectrum*, vol. 38, no. 7, pp. 40–46, July 2001.
12. R. K. Martin, G. Ysebaert and K. Vanbleu, "Bit error rate minimizing channel shortening equalizers for cyclic prefixed systems," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2605–2616, June 2007.
13. "Wikipedia, Orthogonal Frequency Division Multiplexing," <http://en.wikipedia.org/wiki/OFDM>, January 2009.
14. G. K. H. Sari and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, February 1995.

15. D. D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, April 2002.
16. S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice Hall, 2002.
17. D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Systems Technical Journal*, pp. 1541–1562, November 1973.
18. J.A.C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
19. I. Medvedev and V. Tarokh, "A channel shortening multiuser detector for DS-CDMA systems," in *Proc. 53rd Vehicular Tech. Conf.*, vol. 3, 2001, pp. 1834–1838.
20. S. I. Husain and J. Choi, "Single correlator based UWB receiver implementation through channel shortening equalizer," in *Asia-Pasific Conf. Communications*, Perth, Western Australia, October 2005, pp. 610–614.
21. M. Kallinger and A. Mertins, "Room impulse response shortening by channel shortening concepts," in *Conf. Rec. 39th Asilomar Conf. Signals, Syst. Computers*, Pacific Grove, CA, November 2005, pp. 898–902.
22. R. K. Martin, "Blind, adaptive equalization for multicarrier receivers," Ph.D. dissertation, Cornell University, May 2004.
23. G. Arslan, B. L. Evans and S. Kiaei, "Equalization for discrete multitone receivers to maximize bit rate," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3123–3135, December 2001.
24. M. Milošević, L. F. C. Pessoa, B. L. Evans and R. Baldick, "Optimum time domain equalization design for maximizing data rate of discrete multi-tone systems," *IEEE Trans. Signal Process.*, 2003.
25. K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate-maximizing time-domain equalizer design for DMT-based systems," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 871–876, June 2004.
26. P. J. W. Melsa, R. C. Younce and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 1662–1672, December 1996.
27. N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 56–64, January 1996.
28. R. K. Martin *et al.*, "A blind adaptive TEQ for multicarrier systems," *IEEE Signal Process. Lett.*, vol. 9, no. 11, pp. 341–343, November 2002.

29. J. Balakrishnan, R.K.Martin and C.R.Johnson.Jr., "Blind, adaptive channel shortening by sum-squared auto-correlation minimization (SAM)," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3086–3092, December 2003.
30. R.Nawaz and J.A. Chambers, "A novel single lag auto-correlation minimization (SLAM) algorithm for blind adaptive channel shortening," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, March 2005, pp. 885–888.
31. C. Toker and G. Altin, "Blind adaptive channel shortening equalizer algorithm which can provide shortened channel state information (BACS-SI)," *Accepted for publication to Trans. Signal Process.*, 2009.
32. S. Celebi, "Inter block interference (IBI) minimizing time domain equalzier (TEQ) for OFDM," *IEEE Sinnal Process. Lett.*, vol. 10, no. 8, 2003.
33. K. Van Acker *et al.*, "Per tone equalization for DMT-based systems," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 109–119, January 2001.
34. J. S. Chow and J. M. Cioffi, "A cost-effective maximum likelihood receiver for multicarrier systems," in *Proc. IEEE Int. Conf. Commun.*, vol. 2, June 1992, pp. 948–952.
35. N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 903–915, May 1996.
36. N. Al-Dhahir, "FIR channel shortening equalizers for MIMO ISI channels," *IEEE Trans. Commun.*, vol. 49, no. 2, pp. 213–218, February 2001.
37. R. Schur and J. Speidel, "An efficient equalization method to minimize delay spread in OFDM/DMT systems," in *Proc. IEEE Int. Conf. on Commun.*, vol. 5, Helsinki, Finland, pp. 1481–1485.
38. J. G. Proakis, *Digital Communications*, 4th ed., ser. Electrical Engineering Series. New York: McGraw-Hill, Inc., 2001.
39. M. Milošević, L. F. C. Pessoa, and B. L. Evans, "Simultaneous multichannel time domain equalizer design based on the maximum composite shortening SNR," in *Proc. IEEE Asilomar Conf. Sig., Sys. and Comp.*, Pacific Grove, CA.
40. V. Koivunen, J. Laurila, and E. Bonek, *Review of Radio Sicience*. New York: Wiley-IEEE Press, August 2002.
41. R. K. Martin and C. R. Johnson, "Adaptive equalization: Transitioning from single-carrier to multicarrier systems," *IEEE Signal Process. Mag.*, pp. 108–122, Nowember 2005.
42. M. de Courville, P. Duhamel, P. Madec, and J. Palicot, "Blind equalization of OFDM systems based on the minimization of a quadratic criterion," in *Proc. IEEE Int. Conf. Commun.*, June 1996, pp. 1318–1321.

43. T. Pollet, M. Peeters, M. Moonen, and L. Vandendorpe, "Equalization for DMT-based broadband modems," *IEEE Commun. Mag.*, vol. 38, pp. 106–113, May 2000.
44. ITU G.992.1, "Asymmetric Digital Subscriber Line (ADSL) Transceivers," 1999.
45. "ADSL Modem," <http://www.vocal.com/adslmodem.pdf>, January 2009.
46. J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.
47. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. MA: Addison-Wesley, 1989.
48. Randy L. Haupt and Sue E. Haupt, *Practical Genetic Algorithms*, 2nd ed. Hoboken, New Jersey: John Wiley and Sons, Inc., 2004.
49. R. K. Martin, "Fast-converging blind adaptive channel-shortening and frequency-domain equalization," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 102–110, January 2007.
50. D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communications systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, 1980.
51. J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 2, pp. 459–471, 1983.
52. F. Zhang, *The Schur complement and its applications*, ser. Numerical Methods and Algorithms. New York: Springer, 2005, vol. 4.
53. IEEE 802.11, <http://grouper.ieee.org/groups/802/11/>, January 2009.
54. ETSI HIPERLAN2 Standard, <http://www.etsi.org/website/technologies/hiperlan.aspx>, January 2009.
55. G. L. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 1996.
56. T. Starr, M. Cioffi, and P. J. Siverman, *Understanding Digital Subscriber Line Technology*. Prentice Hall PTR, 1999.

Vita

Lieutenant Gökhan Altın graduated from Maltepe Military High School in Izmir, Turkey. He entered undergraduate studies at the Turkish Air Force Academy, Istanbul and graduated with a Bachelor of Science degree in Electronics Engineering in August 2001. After the Academy he attended Communications Officer Training at the Turkish Air Force Comms, Electronics and Inf. Systems School in Izmir, Turkey. Following the Comm. Officer Training, he was assigned to Turkish Air Force Command (TAFC) Comms, Electronics and Inf. Systems Battalion Command as a Platoon Commander in May 2003. In July 2006, he was reassigned to TAFC Personnel Division, Schools and Basic Training Branch as a Project Officer. After serving 1 year, he attended the Graduate School of Engineering and Management, Air Force Institute of Technology in August 2007.

Permanent address: 2950 Hobson Way
Air Force Institute of Technology
Wright-Patterson AFB, OH 45433

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188	
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>						
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE		3. DATES COVERED (From — To)		
10-03-2009		Master's Thesis		Sept 2007 — Mar 2009		
4. TITLE AND SUBTITLE Bit-Error-Rate-Minimizing Channel Shortening using Post-FEQ Diversity Combining and a Genetic Algorithm				5a. CONTRACT NUMBER		
				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Gökhan Altın AFIT, 1Lt, TUAF				5d. PROJECT NUMBER		
				09ENG243		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way WPAFB OH 45433-7765				8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GE/ENG/09-01		
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory 2241 Avionics Circle, Bldg 620 WPAFB, OH 45433 Dr. Vasu Chakravarthy vasu.chakravarthy@wpafb.af.mil DSN:785-5579 x 4245				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/Ryre		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION / AVAILABILITY STATEMENT Approval for public release; distribution is unlimited.						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT Channel shortening filter design is a widely examined topic in the literature. Most of the channel shortening equalizer proposals depend on perfect channel state information (CSI). However, this information may not be available in all situations. In cases where the channel state information is not needed, blind adaptive equalization techniques are appropriate. In wireline communication systems, the CSE design is based on maximizing the bit rate, but in wireless systems, there is a fixed bit loading algorithm, and the performance metric is Bit Error Rate (BER) minimization. In this work, a CSE is developed for multicarrier and single-carrier cyclic prefixed (SCCP) systems which attempt to minimize the BER. To minimize the BER, a Genetic Algorithm (GA) is used. If the CSI is shorter than the CP, the equalization can be done by a frequency domain equalizer (FEQ), which is just a bank of complex scalars. However, in the literature the adaptive FEQ design has not been well examined. The second phase of this thesis focuses on different types of algorithms for adapting the FEQ and modifying the FEQ architecture to get a lower BER.						
15. SUBJECT TERMS adaptive, cyclic prefix, Channel Shortening Equalizer (CSE), Frequency Domain Equalizer (FEQ), Bit Error Rate (BER), Genetic Algorithm (GA)						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Richard K. Martin	
U	U	U	UU	89	19b. TELEPHONE NUMBER (include area code) (937) 255-3636, ext 4625; richard.martin@afit.edu	